Henkilötiedon käsite ja anonyymit tiedot eurooppalaisessa tietosuojalainsäädännössä

Helsingin yliopisto
Oikeustieteellinen tiedekunta
Jesse Heiskanen
Pro gradu -tutkielma
Viestintä- ja informaatio-oikeus
Ohjaajat: Päivi Korpisaari ja
Susanna Lindroos-Hovinheimo
Elokuu 2019

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Tiedekunta – Fakultet – Faculty Oikeustieteellinen tiedekunta		Koulutusohjelma – Utbildingsprogram – Degree Programme Oikeustieteen maisteri
Tekijä – Författare – Author Jesse Heiskanen		
Työn nimi – Arbetets titel – Title Henkilötiedon käsite ja anonyymit tiedot eurooppalaisessa tietosuojalainsäädännössä		
Oppiaine/Opintosuunta – Läroämne/Studieinriktning – Subject/Study track Viestintä- ja informaatio-oikeus		
Työn laji – Arbetets art – Level Pro gradu -tutkielma	Aika – Datum – Month and year Elokuu 2019	Sivumäärä – Sidoantal – Number of pages xxvi + 93

Tiivistelmä – Referat – Abstract

Nykyajalle on leimallista, että ihmisistä kerätään huomattavasti enemmän tietoa kuin koskaan aikaisemmin ja yhä useamman toimijan liiketoiminta perustuu tiedon hyödyntämiseen. Ennennäkemätön tiedonkeruu on johtanut siihen, että anonymiteetti on nyky-yhteiskunnassa entistä vaikeampi saavuttaa, sillä yksilöistä kerätyn tiedon määrän kasvaessa ja teknologian edelleen kehittyessä yksittäisiä tiedonpalasia on mahdollista yhdistää toisiin tietoihin entistä useammin henkilön tunnistamisen mahdollistavalla tavalla. Tällaisessa tilanteessa henkilötietojen suojaa koskevan lainsäädännön merkitys korostuu, kun yhä useammin niin julkiset kuin yksittäiset toimijat käsittelevät toimintansa yhteydessä henkilötietoja. Eurooppalaisen tietosuojalainsäädännön perusta muodostuu Euroopan unionin perusoikeuskirjan henkilötietojen suojaa koskevasta 8 artiklasta, Euroopan unionin toiminnasta tehdyn sopimuksen 16 artiklasta ja näistä ajankohtaisimmasta ja samalla yksityiskohtaisimmasta EU:n yleisestä tietosuoja-asetuksesta.

Tutkielmassa tarkastellaan henkilötiedon käsitteen määritelmää eurooppalaisessa tietosuojalainsäädännössä. Henkilötiedon käsitteen tulkinnalla – sillä, millaiset tiedot katsotaan henkilötiedoiksi – on huomattava merkitys niin julkisille kuin yksityisille toimijoille Euroopassa. Jos käsittelyn kohteena olevien tietojen katsotaan olevan henkilötietoja, tulee niihin soveltaa tietosuojalainsäädännön velvoitteita. Tutkielman tarkoituksena on selvittää, mitä kaikkea henkilötiedon määritelmän alaan eurooppalaisessa tietosuojalainsäädännössä sisältyy. Olennainen osa tätä on sen arvioiminen, mitä kaikkea itse asiassa jää henkilötiedon käsitteen määritelmän ulkopuolelle. Lisäksi kysymys tietosuojalainsäädännön mukaisen henkilötietojen anonymisoinnin mahdollisuudesta, tai toisaalta mahdottomuudesta, liittyy tutkielman aiheeseen olennaisesti.

Henkilötiedon käsitteen määritelmä muodostuu neljästä toisiinsa vahvasti linkittyneestä osatekijästä, jotka ovat ”kaikenlaiset tiedot”, ”liittyvä”, ”tunnistettu tai tunnistettavissa oleva” ja ”luonnollinen henkilö”. Näin ollen henkilötiedon määritelmä on erittäin laaja, muttei kuitenkaan rajaton. Henkilötiedon käsitteen tarkastelu näiden osatekijöiden valossa osoittaa, että kaikenlaiset tiedot voivat olla henkilötietoja ja tiedot voivat liittyä henkilöön tietosuoja-asetuksen edellyttämällä tavalla, kun niiden käsittelyllä on todennäköisesti vaikutuksia henkilön etujen ja oikeuksien kannalta. Henkilötiedon käsitteen ja samalla koko tietosuojalainsäädännön soveltamisalan kannalta merkityksellisin osatekijä on kuitenkin tunnistettavuus, joka useimmiten määrittää sen, ovatko kyseessä olevat tiedot henkilötietoja. Tunnistettavuuden arviointi perustuu kohtuullisen todennäköisesti käytettävissä olevien keinojen arviointiin, joiden alaa sekä tietosuojatyöryhmä että Euroopan unionin tuomioistuin ovat tulkinneet laveasti. Muutenkin viimeaikaista Euroopan unionin tuomioistuimen ratkaisukäytäntöä on leimannut toive varmistaa tehokas henkilötietojen suoja, jonka seurauksena henkilötiedon käsitteen tulkinnankin on nähty laajentuvan koskemaan mitä moninaisempia tietoja. Henkilötiedon käsitteen neljännellä osatekijällä, eli luonnollisella henkilöllä tarkoitetaan tietosuojalainsäädännön kontekstissa sitä, että henkilötietojen suojaa koskevia säännöksiä sovelletaan eläviin henkilöihin heidän kansalaisuudestaan ja asuinpaikastaan riippumatta.

Henkilötietojen anonymisoinnilla tarkoitetaan prosessia, jonka seurauksena henkilötiedoista poistetaan tunnistettavuus siten, ettei rekisteröidyn tunnistaminen ole enää mahdollista. Tietosuojalainsäädännön vaatimukset täyttävän henkilötietojen anonymisoinnin kynnys on eurooppalaisessa tietosuojalainsäädännössä asetettu erittäin korkealle, kun tehokkaasti toteutettu anonymisointi edellyttää sitä, että tietojen tulee pysyä anonyymeinä, vaikka ne yhdistettäisiin toisiin tietoihin, jotka voivat olla käytännössä kenen tahansa hallussa. Henkilötietojen anonymisoinnista voi seurata kuitenkin useita hyötyä rekisterinpitäjinä toimiville, joten useat organisaatiot pyrkivät muodostamaan hallussaan olevista henkilötiedoista anonyymejä tietoja siten, että kyseisiä tietoja olisi mahdollista analysoida vapaasti ilman tietosuojalainsäädännön velvoitteita.

Avainsanat – Nyckelord – Keywords Henkilötietojen suoja, Tietosuoja, Perusoikeudet, EU-oikeus, Henkilötiedon käsite, Anonyymit tiedot, Henkilötietojen anonymisointi, Data protection, Concept of personal data, Anonymization
Ohjaaja tai ohjaajat – Handledare – Supervisor or supervisors Päivi Korpisaari ja Susanna Lindroos-Hovinheimo
Säilytyspaikka – Förvaringställe – Where deposited Helsingin yliopiston kirjasto
Muita tietoja – Övriga uppgifter – Additional information

Sisällys

Lähteet

Kirjallisuus ja artikkelit

Aarnio, Aulis: Oikeussäännösten systematisointi ja tulkinta. Teoksessa *Häyhä, Juha*: Minun metodini. Werner Söderström lakitieto, Helsinki 1997. (Aarnio 1997)

Beck, Gunnar: The Legal Reasoning of the Court of Justice of the EU. Oxford: Hart Publishing 2013. (Beck 2013)

Bengoetxea, Joxerramon: The Legal Reasoning of the European Court of Justice: Towards a European Jurisprudence. Clarendon Press1993. (Bengoetxea 1993)

Beyleveld, Deryck. – Townend, David M.R.: When is personal data rendered anonymous? Interpreting recital 26 of directive 95/46/EC. Medical Law International, 2004, Volume 6, s. 73–86. (Beyleveld – Townend 2004)

Bolognini, Luca – Bistolfi, Camilla: Pseudonymization and impacts of Big (personal/anonymous) Data Processing in the transition from the Directive 95/46/EC to the new EU General Data Protection Regulation. Computer law & Security Review 33 (2017) s. 171–181. (Bolognini – Bistolfi 2017)

Bonnici, Jeanne Pia Mifsud: Exploring the non-absolute nature of the right to data protection. International Review of Law, Computers & Technology, 2014, Volume 28, Issue 2, s.131–143. (Bonnici 2014)

Bottis, Maria – Bouchagiar, George: Personal data v. Big Data in the EU: Control Lost, Discrimina tion Found. Open Journal of Philosophy, 2018, 8, s. 192–205. (Bottis – Bouchagiar 2018)

Bygrave, Lee: Information Concepts in Law: Generic Dreams and Definitional Daylight. Oxford Journal of Legal Studies, Volume 35, Issue 1, 2014, s. 91–120. (Bygrave 2014a)

Bygrave, Lee: Data Privacy Law: An International Perspective. Oxford University Press, 2014. (Bygrave 2014b)

Brynjolfsson, Erik – McAfee, Andrew: The second machine age, work progress and prosperity in a time of brilliant technologies. W.W. Norton & Company 2014. (Brynjolfsson – McAfee 2014)

Cohen, Aloni – Nissim, Kobbi: Towards Formalizing the GDPR's Notion of Singling Out. ArXiv 2019. (Cohen – Nissim 2019)

Conway, Gerard: The Limits of Legal Reasoning and the European Court of Justice. Cambridge University Press 2012. (Conway 2012)

de Hert, Paul — Papakonstantinou, Vagelis — Wright, David — Gutwirth, Serge: The proposed Regulation and the construction of a principles-driven system for individual data protection. Innovation: The European Journal of Social Science Research, Volume 26, Issue 1–2, s. 133–144. (de Hert et al. 2013)

de Hert, Paul — Papakonstantinou, Vagelis: The new General Data Protection Regulation: Still a sound system for the protection of individuals? Computer Law & Security Review, Volume 32, Issue 2, s. 179–194, 2016. (de Hert — Papakonstantinou 2016)

El Khoury, Alessandro: Dynamic IP Addresses Can Be Personal Data, Sometimes. A Story of Binary Relations and Schrödinger's Cat. European Journal of Risk Regulation, 8 (2017), s. 191–197. (El Khoury 2017)

El Emam, Khaled – Álvarez, Cecilia, 2015. A Critical Appraisal of the Article 29 Working Party Opinion 05/2014 on Data Anonymization Techniques. International Data Privacy Law, Volume 5, Issue 1. s. 73–87. (El Emam – Álvarez)

Esayas, Samson Yoseph: The role of anonymisation and pseudonymization under the EU data privacy rules: beyond the "all or nothing" approach. European Journal of Law and Technology, Volume 6, No 2, 2015. (Esayas 2015)

Floridi, Luciano: Is Semantic Information Meaningful Data? Philosophy and Phenomenological Research Volume LXX, No. 2, March 2005. (Floridi 2005)

González Fuster, Gloria: The Emergence of Personal Data Protection as a Fundamental Right of the EU. Springer 2014. (González-Fuster 2014)

González Fuster, Gloria – Gutwirth, Serge: Opening Up Personal Data Protection: A Conceptual Controversy. Computer Law & Security Review, Volume 29, Issue 5, s. 531–539, 2013. (González Fuster – Gutwirth 2013)

Hijmans, Hielke: The European Union as a Guardian of Internet Privacy – The Story of Art 16 TFEU. Springer 2016. (Hijmans 2016)

Hildebrand, Mireille: Profiling and the Identity of the European Citizen. *Teoksessa M. Hildebrandt – S. Gutwirth* (toim.), Profiling the European Citizen: 303 Cross-Disciplinary Perspectives. Springer Science + Business Media B.V. 2008, s. 303–343. (Hildebrandt 2008)

Hintze, Mike: Viewing the GDPR through a de-identification lens: a tool for compliance, clarification, and consistency, International Data Privacy Law, 2018, Volume 8, No. 1. (Hintze 2018)

Hirvonen, Ari: Mitkä metodit? Opas oikeustieteen metodologiaan. Yleisen oikeustieteen julkaisuja 17. Helsinki 2011. (Hirvonen 2011)

Häyhä, Juha: Minun metodini. Weder Söderström lakitieto 1997. (Häyhä 1997)

Jasserand, Catherine: Legal Nature of Biometric Data: From "Generic" Personal Data to Sensitive Data: Which Changes Does the New Data Protection Framework Introduce? European Data Protection Law Review, 2016, Volume 2, Issue 3, s. 297–311. (Jasserand 2016)

Jändel, Magnus: Decision support for releasing anonymised data. Computers & Security, Volume 46, 2014, s. 48–61. (Jändel 2014)

Jääskinen, Niilo: Elements of Style: Julkisasiamiehen ratkaisuehdotus tekstinä. Helsinki law review, Volume 9, Issue 1, 2015, s. 71–73. (Jääskinen 2015)

Kaisto, Janne: Lainoppi ja Oikeusteoria – Oikeusteorian perusteista aineellisen varallisuusoikeuden näkökulmasta. Edita 2007. (Kaisto 2007)

Kindt, Els J: Privacy and Data Protection Issues of Biometric Applications: a Comparative Legal Analysis. Springer 2013. (Kindt 2013)

Kiss, Attila — Szőke, Gergely László: Evolution or Revolution? Steps Forward to a New Generation of Data Protection Regulation. Teoksessa *Gutwirth, Serge — Leenes, Ronald — de Hert, Paul* (toim.): Reforming European Data Protection Law. Law, Governance and Technology Series, Volume 20. Springer, Dordrecht, 2015. (Kiss — Szőke 2015)

Koillinen, Mikael: Henkilötietojen suoja itsenäisenä perusoikeutena. Oikeus 2012 (42); 2, s. 171–193. (Koillinen 2012)

Koillinen, Mikael: Hallinnolliset seuraamukset tietosuojan sanktiomekanismina. Defensor Legis 2016/4, s. 570–586. (Koillinen 2016)

Kokott, Juliane – Sobotta, Christoph: The Distinction Between Privacy and Data Protection in the Jurisprudence of the CJEU and the ECtHR. International Data Privacy Law, 2013, Volume 3, No. 4. (Kokott – Sobotta 2013)

Koops, Bert-Jaap: The Trouble with European Data Protection. International Data Privacy Law, 2014, Volume 4, No. 4, s. 250–261. (Koops 2014)

Korhonen, Rauno: Perusrekisterit ja Tietosuoja. Edita Publishing Oy, 2003. (Korhonen 2003)

Korja, Juhani: Biometrinen tunnistaminen ja henkilötietojen suoja: Tutkimus biometristen tunnisteiden lainsäädännöllisestä asemasta. Lapin yliopisto, Rovaniemi 2016. (Korja 2016)

Korpisaari, Päivi – Pitkänen, Olli – Warma-Lehtinen Eija: Uusi Tietosuojalainsäädäntö. Alma Talent Oy 2018. (Korpisaari et al. 2018)

Korpisaari, Päivi: Oikeudenalan tunnusmerkeistä ja oikeudenalajaotuksen tarpeellisuudesta. Lakimies 7-8/2015, s. 987–1004. (Korpisaari 2015)

Kounadi, Ourania – Resch, Bernd – Petutschnig, Andreas: Privacy threats and protection recommendations for the use of geosocial network data in research. Social Sciences, 2018, 2(10), s. 1–17. (Kounadi et al. 2018)

Leonard, Peter: Customer data analytics: privacy settings for ”big data” business. International Data Privacy Lw, 2014, Volume 4, No. 1. (Leonard 2014)

Li, Dong – He, Xianmang – Cao, LongBin – Chen, Huahui: Permutation anonymization. Journal of Intelligent Information Systems 2016, Volume 47(3), s. 427–445. (Li et al. 2016)

Lindroos-Hovinheimo: Henkilötietojen suoja perusoikeutena – yksityisyyttä yhteisön kustannuksella? Lakimies 1/2018, s. 52–75. (Lindroos-Hovinheimo 2018)

Lock, Tobias: The European Court of Justice and International Courts. Oxford University University 2015. (Lock 2015)

Lynskey, Orla: The Foundations of EU Data Protection Law. Oxford Studies in European Law 2015. (Lynskey 2015)

Lynskey, Orla: Deconstructing Data Protection: The 'Added-Value' of a Right to Data Protection in the EU Legal Order. International and Comparative Law Quarterly, 63 (3), s. 569–597. (Lynskey 2014)

Mayer-Schönberger, Viktor — Cukier, Kenneth: Big Data. A revolution that will transform how we live, work and think. John Murray, 2013. (Mayer-Schönberger — Cukier 2013)

Mayer-Schönberger, Viktor – Padova, Yann: Regime Change? Enabling Big Data Through Europe’s New Data Protection Regulation. The Columbia Science & Technology Law Review, Colum. Sci. & Tech. L. Rev., Volume 17, s. 315–402, 2016. (Mayer-Schönberger – Padova 2016)

Mittelstadt, Brent – Wachter, Sandra: a right to reasonable inferences: re-thinking data protection law in the age of big data and ai. Columbia Business Law Review, 2019(2). (Mittelstadt – Wachter 2019)

Mostert, Menno – Bredenoord, Annelien L – Biesaart, Monique CIH, van Delden, Johannes JM: Big Data in medical research and EU data protection law: challenges to the consent or anonymise approach. 2015. (Mostert et al. 2015)

Mourby, Miranda – Mackey, Elaine – Elliot, Mark – Gowans, Heather – Wallace, Susan E. – Bell, Jessica – Smith, Hannah – Aidinlis – Stergios – Kaye, Jane: Are ”pseudonymised” data always personal data? Implications of the GDPR for administrative data research in the UK. Computer Law & Security Review 34, s. 222–233, 2018. (Mourby et al. 2018)

Mäenpää, Olli: Eurooppalainen hallinto-oikeus. Talentum 2011. Kolmas uudistettu painos. (Mäenpää 2011)

Mäenpää, Olli: Julkisuusperiaate. Talentun Pro, Helsinki 2016. (Mäenpää 2016)

Narayanan, Arvind – Vitaly Shmatikov: Robust De-anonymization of Large Sparse Datasets. 2008 IEEE Symposium on Security and Privacy, s. 111–125. (Narayan – Shmatikov 2008)

Sethi, Nayha – Laurie, Grame T.: Delivering proportionate governance in the era of eHealth: Making linkage and privacy work together. Medical Law International 2013, Volume 12, Issue 2–3, s. 168–204. (Sethi – Laurie 2013)

Neuvonen, Riku: Viestintä- ja informaatio-oikeuden perusteet. Toinen painos, Kauppakamari, Helsinki 2019. (Neuvonen 2019)

Nissenbaum, Helen: The Meaning of Anonymity in an Information Age. The Information Society: An International Journal, Volume 15, 1999, Issue 2, s. 141–144. (Nissenbaum 1999)

Elliot, Mark – O'Hara, Kieron – Raab, Charles – O'Keefe, Christine M. – Mackey, Elaine – Dibben, Chris – Gowans, Heather – Purdam, Kingsley – McCullagh, Karen: 'Functional anonymisation: Personal data and the data environment' Computer Law & Security Review, Volume 34, no. 2, s. 204–221. (Elliot et al. 2018)

Ohm, Paul: Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. UCLA Law Review, Volume 57, 2010(2009), s. 1701–1819. (Ohm 2009)

Ojanen, Tuomas: EU-oikeuden perusteita. Edita 2016. (Ojanen 2016)

Omer, Tene – Polonetsky, Jules: Big Data for All: Privacy and User Control in the Age of Analytics 2013. (Omer – Polonetsky 2013)

Oswald, Marion: Share and share alike? An examination of trust, anonymization and data sharing with particular reference to an exploratory research project investigating attitudes to sharing personal data with the public sector 2014. (Oswald 2014)

Pabst, Sabine: Unbeobachtete Kommunikation. Das Konzept von Anonymität im Mediendiskurs seit der Aufklärung. Springer Vs. 2018. (Pabst 2018)

Paunio, Elina – Lindroos-Hovinheimo, Susanna: Kielellisen tulkinnan fiktio EU-oikeudessa. Lakimies 2/2008, s. 230–247. (Paunio – Lindroos-Hovinheimo 2008)

Paunio, Elina – Lindroos-Hovinheimo, Susanna: Taking Language Seriously: An Analysis of Linguistic Reasoning and Its Implications in EU Law. European Law Journal July 2010, Volume 16, Issue 4. Sivut 395–416. (Paunio – Lindroos-Hovinheimo 2010)

Paunio, Elina: Legal Certainty in Multilingual EU Law: Language, Discourse and Reasoning at the European Court of Justice. Ashgate 2013. (Paunio 2013)

Pellonpää, Matti – Gullans, Monica – Pölönen, Pasi – Tapanila, Antti: Euroopan ihmisoikeussopimus. 6. uudistettu painos. Alma Talent Oy 2018. (Pellonpää et al. 2018)

Petkova, Bilyana: Privacy as Europe's First Amendment. European Law Journal, Volume 25, Issue 2, 2019, s. 140–154. (Petkova 2019)

Pitkänen – Tiilikka – Warma: Henkilötietojen suoja. Talentum 2014. (Pitkänen et al. 2014).

Poullet, Yves: Is the general data protection regulation the solution? Computer Law &Security Review, 34(4), 2018, s. 773–778. (Poullet 2018)

Purtova, Nadezhda: Default entitlements in personal data in proposed Regulation: Informational self-determination off the table…and back on again? Computer Law & Security Review 30 (2014) s. 6–24. (Purtova 2014)

Purtova, Nadezhda: The law of everything. Broad concept of personal data and future of EU data protection law. Law, Innovation and Technology, 2018, Volume 10(1), s. 40–81. (Purtova 2018)

Quinn, Paul: The anonymization of research data – a pyric victory for privacy that should not be pushed too hard by the EU data protection framework? European Journal of Health Law, 24(4), 2017, s. 347–367. (Quinn 2017)

Quinn, Paul – Quinn, Liam: Big Genetic Data and Its Big Data Protection Challenges. Computer Law & Security Review 34 (2018) s. 1000–1018. (Quinn – Quinn 2018)

Quelle, Claudia: The "Risk Revolution" in EU Data Protection Law: We can't Have Our Cake and Eat it, Too. Teoksessa: *Leenes, Ronald – van Brakel, Rosamunde – Gutwirth, Serge – De Hert, Paul: Data Protection and Privacy: The Age of Intelligent Machines.* Hart Publishing 2018. (Quelle 2018)

Raitio, Juha: Euroopan unionin oikeus. Talentum Pro, Helsinki 2016. (Raitio 2016)
Raitio, Juha: Oikeusvaltion ääriviivat. Alma Talent 2017. (Raitio 2017)

Reding, Viviana: The European data protection framework for the twenty-first century. International Data Privacy Law, Volume 2. Issue 3, 2012, s. 119–129. (Reding 2012)

Rosas, Allan: Perus- ja ihmisoikeudet EU-oikeudessa. Teoksessa: *Hallberg, Pekka – Karapuu, Heikki – Ojanen, Tuomas – Scheinin, Martin – Tuori, Kaarlo – Viljanen, Veli-Pekka*: Perusoikeudet 2011. WSOY Pro, Oikeuden Perusteokset. (Rosas 2011)

Sankari, Suvi: European Court of Justice Legal Reasoning in Context. Europa Law Publishing. Groningen 2013. (Sankari 2013)

Schwartz, Paul M. – Solove, Daniel J: The PII problem: privacy and a new concept of personally identifiable information. New York University Law Review 86(6), 2011, s. 1814–1894. (Schwartz – Solove 2011)

Schwartz, Paul M. – Solove, Daniel J: 'Reconciling Personal Information in the United States and European Union. UC Berkeley Public Law Research Paper No. 2271442. (Schwartz – Solove 2014)

Shannon, Claude: A Mathematical Theory of Communication. The Bell System Technical Journal, Volume 27, s. 379–423, 623–656, July, October 1948. (Shannon 1948)

Siltala, Raimo: Oikeustieteen tieteenteoria. Helsinki 2003. (Siltala 2003)

Solove, Daniel J: Understanding Privacy. Harvard University Press, Cambridge, Massachusetts; London, England. (Solove 2008)

Stalla-Bourdillon, Sophie – Knight, Alison: Anonymous data v. Personal data – a false debate: An EU perspective on anonymisation, pseudonymisation and personal data. 2016. (Stalla-Bourdillon – Knight 2016)

Stevens, L.: The proposed data protection regulation and its potential impact on social sciences research in the UK. European Data Protection Law Review, Volume 1, no. 2, 2015, s. 97–122. (Stevens 2015)

Sweeney, Latanya: k-Anonymity: A Model for Protecting Privacy. International Journal on Uncertainty, Fuzziness and Knowledge-based Systems 10(5) 2002, s. 557–570. (Sweeney 2002)

Tamò Larrieux, Aurelia: Designing for Privacy and its Legal Framework: Data Protection by Design and Default for the Internet of Things. Springer 2018. (Tamò Larrieux 2018)

Tarhonen, Laura: Pseudonymisation of personal data according to the general data protection regulation. Edilex 2017. (Tarhonen 2017)

Tuori, Kaarlo: Kriittinen Oikeuspositivismi. Werner Söderström lakitieto, Helsinki 2000. (Tuori 2000)

Urgessa, Worku Gedefa: The Protective Capacity of the Criterion of 'Identifiability' under EU Data Protection Law. European Data Protection Law Review, Volume 2, 2016, Issue 4, s. 521–531. (Urgessa 2016)

Van Alsenoy, Brendan: Data Protection Law in the EU: Roles, Responsibilities and Liability. Intersentia Ltd, Cambridge 2019. (Van Alsenoy 2019)

van der Sloot, Bart: Do data protection rules protect the individual and should they? An assessment of the proposed General Data Protection Regulation. International Privacy Law, 2014, Volume 4, No. 4. (van der Sloot 2014)

van Loenen, Bastiaan – Kulk, Stefan – Ploeger, Hendrik: Data protection legislation: A very hungry caterpillar The case of mapping data in the European Union. Government Information Quarterly: an International Journal of Information Technology Management, Policies, and Practices, Volume 33, No. 2, 2016, s. 338–345. (van Loenen et al. 2016)

Wallace, Kathleen A.: Anonymity. Ethics and Information Technology. Volume 1, s. 21–31. 1999. (Wallace 1999)

Warren, Samuel D, – Brandeis, Louis D: The Right to Privacy. Harward Law Review, Volume 4, No. 5. (Dec. 15, 1890), s. 193–220. (Warren – Brandeis 1980)

Yue, Liu: Identifying Legal Concerns in the Biometric Context. Journal of International Commercial Law and Technology. 2008, Volume 3, Issue 1. (Yue Liu 2008)

Zuiderveen Borgesius, Frederik J. – Gray, Jonathan – van Eechoud, Mireille: Open Data, Privacy, and Fair Information Principles: Towards a Balancing Framework. Berkeley Technology Law Journal, 2015, Volume 30(3), s. 2073–2131. (Zuiderveen Borgesius et al. 2015)

Zuiderveen Borgesius, Frederik J.: Singling out people without knowing their names – Behavioural targeting, pseudonymous data, and the new Data Protection Regulation. Computer Law & Security Review, Volume 32, Issue 2, 2016, s. 256–271. (Zuiderveen Borgesius 2016)

Zuiderveen Borgesius, Frederik J.: Breyer Case of the Court of Justice of the European Union: IP Addresses and the Personal Data Definition. European Data Protection Law Review, Volume 3, No. 1, 2017, s. 130–137. (Zuiderveen Borgesius 2017)

Virallislähteet

Kansainväliset sopimukset

Euroopan neuvoston yleissopimus ihmisoikeuksien ja perusvapauksien suojaamiseksi (4.11.1950, SopS 63/1999). (Euroopan ihmisoikeussopimus, EIS)

Euroopan neuvoston yleissopimus yksilöiden suojelusta henkilötietojen automaattisessa tietojenkäsittelyssä. (28.1.1981, ETS nro 108, SopS 36/1992) (Tietosuojasopimus)

Euroopan unionin primaarioikeus

Euroopan unionin perusoikeuskirjan konsolidoitu toisinto, EUVL C 326, 26.10.2012, s. 391–407. (Euroopan unionin perusoikeuskirja, POK)

Euroopan unionin toiminnasta tehdyn sopimuksen konsolidoitu toisinto, EUVL C 326, 26.10.2012, s. 47–390. (Sopimus Euroopan unionin toiminnasta, SEUT)

Euroopan unionista tehdyn sopimuksen konsolidoitu toisinto, EUVL C 326, 26.10.2012, s. 13–46. (Sopimus Euroopan unionista, SEU)

Lissabonin sopimus Euroopan unionista tehdyn sopimuksen ja Euroopan yhteisön perustamissopimuksen muuttamisesta, EUVL N:o C 306, 17.12.2007, s. 1. (Lissabonin sopimus)

Euroopan unionin sekundaarioikeus

Euroopan parlamentin ja neuvoston direktiivi 95/46/EY, annettu 24 päivänä lokakuuta 1995, yksilöiden suojelusta henkilötietojen käsittelyssä ja näiden tietojen vapaasta liikkuvuudesta. EYVL N:o L 281, 23.11.1995, s. 31–50. (Henkilötietodirektiivi)

Euroopan parlamentin ja neuvoston direktiivi 2002/58/EY, annettu 12 päivänä heinäkuuta 2002, henkilötietojen käsittelystä ja yksityisyyden suojasta sähköisen viestinnän alalla. (Sähköisen viestinnän tietosuojadirektiivi)

Euroopan parlamentin ja neuvoston direktiivi 2003/98/EY, annettu 17 päivänä marraskuuta 2003, julkisen sektorin hallussa olevien tietojen uudelleenkäytöstä. (PSI-direktiivi)

Euroopan parlamentin ja neuvoston direktiivi 2006/24/EY, annettu 15 päivänä maaliskuuta 2006, yleisesti saatavilla olevien sähköisten viestintäpalvelujen tai yleisten viestintäverkkojen yhteydessä tuotettavien tai käsiteltävien tietojen säilyttämisestä ja direktiivin 2002/58/EY muuttamisesta. (Direktiivi 2006/24/EY)

Euroopan parlamentin ja neuvoston asetus (EU) 2016/679, annettu 27 päivänä huhtikuuta 2016, luonnollisten henkilöiden suojelusta henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta ja direktiivin 95/46/EY kumoamisesta. EUVL N:o L 119, 4.5.2016, s. 1–88. (Yleinen tietosuoja-asetus, TSA)

Euroopan unionin jäsenvaltioiden kansallinen lainsäädäntö

Yhdistynyt kuningaskunta: Data Protection Act 2018. (Data Protection Act 2018)

Tanska: LOV nr 502 af 23/05/2018, Lov om supplerende bestemmelser til forordning om beskyttelse af fysiske personer i forbindelse med behandling af personoplysninger og om fri udveksling af sådanne oplysninger. (Tanskan tietosuojalaki)

Euroopan unionin jäsenvaltioiden virallislähteet

HE 159/17 vp. Hallituksen esitys eduskunnalle laiksi sosiaali- ja terveystietojen toissijaisesti käytöstä sekä eräiksi siihen liittyviksi laeiksi. (HE 159/17 vp.)

Koski, Heli – Honkanen, Mika – Luukkonen, Juha – Pajarinen, Mika – Ropponen, Teemu: Avoimen datan hyödyntäminen ja vaikuttavuus. Valtioneuvoston tutkimustoiminnan julkaisusarja 40/2017. (Koski et al.)

HE 159/2017 vp StV 25.4.2018 tietosuojavaltuutettu Reijo Aarnion asiantuntijalausunto. (Tietosuojavaltuutettu 25.4.2018)

Tietosuojavaltuutetun toimiston toimintakertomus 2018. (TSV 2018)

Ympäristöministeriön raportteja 10 | 2018 Henkilötiedot ja paikkatiedot. *Korpisaari, Päivi:* Miten tietosuojalainsäädäntö vaikuttaa paikkatietojen julkaisemiseen ja luovuttamiseen. (Korpisaari 2018)

Pääministeri Antti Rinteen hallituksen ohjelma 6.6.2019. Osallistava ja Osaava Suomi – sosiaalisesti, taloudellisesti ja ekologisesti kestävä yhteiskunta. (Pääministeri Rinteen hallitusohjelma 2019)

Valtiovarainministeriö: AuroraAI – kohti ihmiskeskeistä yhteiskuntaa. Kansallisen tekoälyohjelma Auroran esiselvityshankkeessa tuotettu kehittämis- ja toimeenpanosuunnitelma 2019–2023. (Valtiovarainministeriö 2019)

Muu Euroopan unionin virallismateriaali

Commission communication on the protection of individuals in relation to the processing of personal data in the Community and information security. COM (90) 314 final, 13.9.1990. (Euroopan komissio 1990)

Amended proposal for a council directive on the protection of individuals with regard to the processing of personal data and on the free movement of such data. COM (92) 422 final, 28.10.1992. (Euroopan komissio 1992)

Europe 2020 – A European strategy for smart, sustainable and inclusive growth. Annettu 3.3.2010. (Europe 2020 strategy)

Euroopan digitaalistrategia. Kattava lähestymistapa henkilötietojen suojaan Euroopan unionissa. Bryssel 19.5.2010 (KOM(2010) 245 lopullinen). (Euroopan komissio 2010)

Komission ehdotus Euroopan parlamentin ja neuvostona asetukseksi yksilöiden suojelusta henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta (yleinen tietosuoja-asetus) Bryssel 25.1.2012 COM(2012) 11 final – 2012/0011 (COD). (Euroopan komissio 2012)

Information Commissioner's Office: Anonymisation: managing data protection risk code of practice. Julkaistu marraskuussa 2012. (ICO 2012)

Komission tiedonanto Euroopan parlamentille, neuvostolle, Euroopan talous- ja sosiaalikomitealle ja alueiden komitealle. Digitaalisten sisämarkkinoiden strategia Euroopalle. Bryssel 6.5.2015. (COM(2015) 192 final. (Digital single market 2015)

European Union Agency for Network and Information Security (ENISA): Privacy by design in big data – An overview of privacy enhancing technologies in the era of big data analytics 2015. (ENISA 2015)

Information Commissioner's Office: ICO analysis of the Council of the European Union text of the General Data Protection Regulation. (ICO 2017a)

Information Commissioner's Office: Big data, artificial intelligence, machine learning and data protection. Julkaistu syyskuussa 2017. (ICO 2017b)

Komission tiedonanto Euroopan parlamentille, Eurooppa-neuvostolle, neuvostolle, Euroopan talous- ja sosiaalikomitealle ja alueiden komitealle. Luotettavat digitaaliset sisämarkkinat kaikille – valmiiksi saattaminen. COM(2018) 320 final. (Euroopan komissio 2018)

Handbook on European data protection law. Publications of the European Unioni, Luxembourg 2018. (Data protection handbook 2018)

European Data Protection Board (EDPB) Work Program 2019–2020. (EDPB Work Program 2019–2020)

Proposal for a Regulation of the European Parliament and of the Council concerning the respect for private life and protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications). Annettu 26.7.2019. (ePrivacy-asetus heinäkuun 2019 luonnos)

Communication from the Commission to the European parliament and the Council. Data Protection rules as a trust-enabler in the EU and beyond – taking stock. Brussels, 24.7.2019. COM(2019 374 final. (Euroopan komissio 2019)

WP 29: Opinion 4/2007 on the concept of personal data. Annettu 20. kesäkuuta 2007. (WP 136)

WP 29: Lausunto 1/2008 hakukoneisiin liittyvistä tietosuojakysymyksistä. Annettu 4. huhtikuuta 2008. (WP 148)

WP 29: Opinion 1/2010 on the concepts of "controller" and "processor". Annettu 16. helmikuuta 2010. (WP 169)

WP 29: Opinion 16/2011 on EASA/IAB Best Practice Recommendation on Online Behavioural Advertising. WP 188. Annettu 8.12.2011. (WP 188)

WP 29: Opinion 05/2014 on Anonymisation Techniques. Annettu 10. huhtikuuta 2014. (WP 216)

WP 29: Opinion 06/2014 on the notion of legitimate interest of the data controller under Article 7 of Directive 95/46/EC. Annettu 9. huhtikuuta 2014. (WP 217)

WP 29: Guidelines for identifying a controller of processor's lead supervisory authority. Annettu 13.12.2016. (WP 244 rev.01)

WP 29: Guidelines on Data Protection Impact assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679. Annettu 4. huhtikuuta 2017. (WP 248)

WP 29: Guidelines on consent under Regulation 2016/679. Annettu 28. marraskuuta 2017. Päivitetty 10. huhtikuuta 2018. (WP 259 rev.01)

OECD

OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data, 23.9.1980. (OECD 1981)

OECD Updated Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. Päivitetty versio, OECD Privacy Framework – The Revised OECD Privacy Guidelines. (OECD 2013)

Oikeuskäytäntö ja muut ratkaisut

Euroopan ihmisoikeustuomioistuin

Asia S and Marper v UK. Tuomio annettu 4.12.2008.

Asia López Ribalda ym. v. Espanja. Tuomio annettu 9.1.2018.

Asia Antovic ja Mirkovic v. Montenegro. Tuomio annettu 28.11.2017

Asia Vukota-Bojic v. Sveitsi. Tuomio annettu 18.10.2016.

Euroopan unionin tuomioistuin

Asia C-283/81 Srl CILFIT ja Lanificio di Gavardo SpA v. Ministero della Sanita. Tuomio annettu 6.10.1982. (C-283/81 CILFIT)

Yhdistetyt asiat C-465/00, C-138/01 ja C-139/01 Rechnungshof v. Österreichischer Rundfunk ym. Tuomio annettu 20.5.2003. (C-465/00 Österreichischer Rundfunk ym.)

Asia C–101/01 Rikosoikeudenkäynti v. Bodil Lindqvist. Tuomio annettu 6.11.2003. (C-101/01 Lindqvist)

Asia C-275/06 Productores de Música de España (Promusicae) v. Telefonica de Espana SAU. Tuomio annettu 29.1.2008. (C-275/06 Promusicae)

Asia C-73/07 Tietosuojavaltuutettu v. Satakunnan Markkinapörssi Oy ja Satamedia Oy. Tuomio annettu 16.12.2008. (C-73/07 Satakunnan markkinapörssi ja Satamedia Oy)

Asia C-553/07 College van burgemeester en wethouders van Rotterdam v. E. E. Rijkeboer. Tuomio annettu 7.5.2009. (C-553/07 Rijkeboer)

Asia C-28/08 P Euroopan komissio v. Bavarian Lager Co. Ltd. Tuomio annettu 29.6.2010. (C-28/08 P Bavarian Lager)

Asia C-280/08 Deutsche Telekom v. Euroopan komissio. Tuomio annettu 14.10.2010. (C-280/08 Deutsche Telekom)

Yhdistetyt asiat C-92/09 ja C-93/09 Volker und Markus Schecke GbR ja Hartmut Eifert (C-93/09) v. Land Hessen. Tuomio annettu 9.11.2010. (C-92/09 Volker und Markus Schecke ja Eifert)

Asia C–70/10 Scarlet Extended SA v. Societe belge des auteurs, compositeurs et editeurs SCRL (SABAM). Tuomio annettu 24.11.2011. (C-70/10 Scarlet Extended)

Asia C-342/12 Worten Equipamentos para o Lar SA v. Autoridade pasa as Condicoes de Trabalho (ACT). Tuomio annettu 30.5.2013. (C-342/12 Worten)

Yhdistetyt asiat C-293/12 ja C-594/12 Digital Rights Ireland Ltd. v. Minister for Communications, Marine and Natural Resources ym. ja Karntner Landesregierung ym. Tuomio annettu 8.4.2014. (C-293/12 Digital Rights Ireland)

Asia C-131/12 Google Spain SL ja Google Inc. v. Agencia Española de Protección de Datos (AEPD) ja Mario Costeja González. Tuomio annettu 13.5.2014. (C-131/12 Google Spain)

Yhdistetyt asiat C-141/12 ja C-372/12 YS. v. Minister voor Immigratie, Integratie en Asiel ja Minister voor Immigratie, Integratie en Asiel v. M ja S. Tuomio annettu 17.7.2015. (C-141/12 YS ym.)

Asia C–362/14 Maximillian Schrems v. Data Protection Commissioner. Tuomio annettu 6.10.2015. (C-362/14 Schrems)

Asia C–582/14 Patrick Breyer v. Saksan liittotasavalta. Tuomio annettu 19.10.2016. (C-582/14 Breyer)

Asia C-434/16 Peter Nowak v. Data Protection Commissioner. Tuomio annettu 20.12.2017. (C-434/16 Nowak)

Asia C-210/16 Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein v. Wirtschaftsakademie Schleswig-Holstein GmbH. Tuomio annettu 5.6.2018. (C-210/16 Wirtschaftsakademie)

Asia C-25/17 Jehovan todistajat v. Tietosuojavaltuutettu. Tuomio annettu 6.11.2018. (C-25/17 Jehovan todistajat)

Euroopan unionin tuomioistuimen julkisasiamiehen ratkaisuehdotukset

Julkiasiamies Cruz Villalónin ratkaisuehdotus asiassa C-70/10 Scarlet Extended. Annettu 14.4.2011.

Julkisasiamies Jääskisen ratkaisuehdotus asiaan C-131/12 Google Spain. Annettu 25.6.2013.

Julkisasiamies Sharpstonin ratkaisuehdotus asiassa C-141/12 YS ym. Annettu 12.12.2013.

Julkisasiamies Sánchez-Bordonan ratkaisuehdotus asiassa C-582/14 Breyer. Annettu 12.5.2016.

Julkisasiamies Kokottin ratkaisuehdotus asiassa C-434/16 Nowak. Annettu 20.7.2017.

Julkisasiamies Bobekin ratkaisuehdotus asiassa C-40/17 FashionID. Annettu 19.12.2018.

Viranomaisratkaisut

Itävallan tietosuojaviranomainen: Datenschutzbehörde, DSB-D123.270/0009-DSB/2018. Annettu 5.12.2018. (DSB 2018)

Tanskan tietosuojaviranomainen: Datatilsynet, DPA 2018-41-0016. Annettu 18.3.2019. (Datatilsynet 2019)

Internet-lähteet

BBC: Fitness app Strava lights up stadd at military bases. Saatavilla: https://www.bbc.com/news/technology-42853072.Vierailtu 6.8.2019. (BBC 29.1.2018)

Personal Data Protection Commission Singapore – Anonymisation: Managing Personal Data Protection Risk. Saatavilla: https://www.pdpc.gov.sg/-/media/Files/PDPC/New_DPO_Connect/nov_15/pdf/Anonymisation.pdf. Vierailtu 1.6.2019. (PDPC 2018)

The Guardian: New York taxi details can be extracted from anonymized data, researchers say. Saatavilla: https://www.theguardian.com/technology/2014/jun/27/new-york-taxi-details-anonymised-data-researchers-warn. Vierailtu 7.8.2019 (Guardian 27.6.2014)

The Guardian: ”Data is a fingerprint”: Why you aren’t as anonymous as you think online. Saatavilla: https://www.theguardian.com/world/2018/jul/13/anonymous-browsing-data-medical-records-identity-privacy. Vierailtu 7.8.2019. (Guardian 13.7.2018)

IEEE Xplore: Anonymizing NYC Taxi Data: Does It Matter? Saatavilla: https://ieeexplore.ieee.org/document/7796899.Vierailtu 4.8.2019. (IEEE Explore 19.10.2016)

Data Protection Commission (DPC) Ireland: Anonymisation and pseudonymisation. Saatavilla: https://www.dataprotection.ie/en/guidance-landing/anonymisation-and-pseudonymisation. Vierailtu 25.5.2019. (DPC 2019)

Technology Science: Only You, Your Doctor, and Many Others May Know. Saatavilla: https://techscience.org/a/2015092903/. Vierailtu 6.8.2019. (Techscience 29.9.2015)

UKAN Publications: The Anonymisation Decision-Making Framework. Saatavilla: https://ukanon.net/wp-content/uploads/2015/05/The-Anonymisation-Decision-making-Framework.pdf. Viitattu 10.8.2019. (UKAN 2016)

Transport for London: Wi-Fi data collection. Saatavilla https://tfl.gov.uk/corporate/privacy-and-cookies/wi-fi-data-collection. Vierailtu: 9.8.2019. (TfL heinäkuu 2019)

Wired: Twitter's vast metadata haul is a privacy nightmare for users. Saatavilla: https://www.wired.co.uk/article/twitter-metadata-user-privacy. Vierailtu 5.8.2019. (Wired 9.7.2018)

Wired: TfL is going to tack all London Underground users using Wi-Fi. Saatavilla: https://www.wired.co.uk/article/london-underground-wifi-tracking. Vierailtu 8.8.2019. (Wired 22.5.2019)

BBC: Chinese man caught by facial recognition at pop concert. Saatavilla: https://www.bbc.com/news/world-asia-china-43751276. Vierailtu 25.6.2019. (BBC 13.4.2018)

The New York Times: San Francisco Bans Facial Recognition Technology. Saatavilla: https://www.nytimes.com/2019/05/14/us/facial-recognition-ban-san-francisco.html. Vierailtu 1.7.2019. (New York Times 14.5.2019)

Sophos: You don't have to sequence your DNA to be identifiable by your DNA. Saatavilla: https://nakedsecurity.sophos.com/2018/10/18/you-dont-have-to-sequence-your-dna-to-be-identifiable-by-your-dna/. Vierailtu 13.6.2019. (Sophos 18.10.2018)

Helsingin sanomat: Tekoäly voi ennustaa lastensuojelun asiakkuuden etukäteen, uskoo Espoo – Tieto tutkii asukkaiden terveystietoja. Saatavilla: https://www.hs.fi/kaupunki/art-2000005709748.html. Vierailtu 10.8.2019. (HS 7.6.2018)

IAPP: Does anonymization or de-identification require consent under the GDPR? Saatavilla: https://iapp.org/news/a/does-anonymization-or-de-identification-require-consent-under-the-gdpr/. Vierailtu 13.7.2019. (IAPP 2019)

Reuters: TomTom cleared of data probe violations. Saatavilla: https://www.reuters.com/article/us-tomtom-idUSTRE80B0HN20120112. Vierailtu 6.6.2019. (Reuters 12.1.2012.)

The New York Times: A Face Is Exposed for AOL Searcher No. 4417749. Saatavilla: https://www.nytimes.com/2006/08/09/technology/09aol.html. Vierailtu 10.8.2019. (New York Times 9.8.2006)

Fast Company: NYC Taxi Data Blunder Reveals Which Celebs Don't Tip – And Who Frequents Strip Clubs. Saatavilla: https://www.fastcompany.com/3036573/nyc-taxi-data-blunder-reveals-which-celebs-dont-tip-and-who-frequents-strip-clubs. Vierailtu 19.8.2019. (Fast Company 10.2.2014)

Lyhenteet

EDPB	European Data Protection Board
EU	Euroopan unioni
EUT	Euroopan unionin tuomioistuin
EIS	Euroopan ihmisoikeussopimus
EIT	Euroopan ihmisoikeustuomioistuin
OECD	Taloudellisen yhteistyön ja kehityksen järjestö (Organization for Economic Co-operation and Development)
SEU	Sopimus Euroopan unionista
SEUT	Sopimus Euroopan unionin toiminnasta
TSA	Euroopan parlamentin ja neuvoston asetus (EU) 2016/679 luonnollisten henkilöiden suojelusta henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta ja direktiivin 95/46/EY kumoamisesta
WP 29	Article 29 Working Party, EU:n tietosuojatyöryhmä

Kaaviot

1 Johdanto

1.1 Tutkielman tausta

Termi anonyymi on peräisin kreikan kielestä, jossa se tarkoittaa nimettömyyttä tai nimeltään tuntematonta. Anonymiteetin konseptilla on kuitenkin nykyaikana huomattavasti laajempi merkitys kuin alkuperäisellä nimettömyyttä kuvaavalla adjektiivilla. Anonyymius tarkoittaa sitä, että henkilö ei ole tunnistettavissa, tavoitettavissa tai jäljitettävissä.[1] Näin ollen anonymiteetti on moniulotteisempi käsite kuin pelkkä nimettömyys ja se muodostuu yksilön suhteesta ympäröivään yhteiskuntaan; toisin sanoen anonymiteetti edellyttää sitä, että tiedämme tietystä henkilöstä jotain, mutta emme tästä huolimatta pysty selvittämään hänen henkilöllisyyttään.[2] Tästä lähtökohdasta anonymiteetin vähimmäisedellytyksenä on mahdollista pitää sitä, että tiedämme jonkin henkilön vähintään olevan olemassa.

Nykyajalle on leimallista, että ihmisistä kerätään huomattavasti enemmän tietoa kuin koskaan aikaisemmin ja yhä useamman toimijan liiketoiminta perustuu tiedon hyödyntämiseen. Tiedonkeruu on viime vuosikymmenten aikana kasvanut räjähdysmäisesti, sillä tieto- ja viestintäteknologia on tullut yhä kiinteämmäksi osaksi päivittäistä elämäämme, ja kerättyä tietoa voidaan nykyaikasilla menetelmillä käsitellä, säilyttää ja analysoida ennennäkemättömillä tavoilla.[3] Tämä kehitys johtaa siihen, että anonymiteetti on nyky-yhteiskunnassa entistä vaikeampi saavuttaa, sillä yksilöistä kerätyn tiedon määrän kasvaessa ja teknologian edelleen kehittyessä yksittäisiä tiedonpalasia on mahdollista yhdistää toisiin tietoihin entistä useammin henkilön tunnistamisen mahdollistavalla tavalla.[4] Tällaiseen kehitykseen liittyvänä uhkakuvana on nähty jopa orwellilaisen dystopian syntyminen, jossa yksilön jokainen viesti ja reaktio tallennetaan profilointitarkoituksessa kafkamaisin seurauksin.[5]

[1] Wallace 1999, s. 24–25; Nissenbaum 1999, s. 141.
[2] Pabst 2018, s. 27–29.
[3] Bottis – Bouchagiar 2018, s. 193–194; Brynjolfsson – McAfee 2014, s. 98–99.
[4] Vaihtoehtoisesti asiaa on mahdollista tarkastella siitä näkökulmasta, että nykyajalle leimallisen massiivisen tiedonkeruun seurauksena yksittäisen henkilön anonymiteetti on yhä useammin mahdollista murtaa. Anonymiteetin merkityksestä laajemmin ks. esim. Solove 2008, s. 125.
[5] Mayer-Schönberger – Cukier 2013, s. 150–151. Kirjailija George Orwell visioi vuonna 1949 julkaistussa kirjassaan ”Vuonna 1984” kansalaisten jatkuvaan valvontaan perustuvan dystopian, josta termi ”orwellilainen” on johdettu. Termi ”kafkamainen” on taas peräisin kirjailija Franz Kafkan tuotannosta, erityisesti hänen vuonna 1925 julkaisusta kirjastaan ”Oikeusjuttu”, jossa päähenkilö asetetaan syytteeseen rikoksesta, jota ei koskaan paljasteta. Termiä käytetään yleisesti kuvaamaan yksilön joutumista kasvottoman valtakoneiston uhriksi.

Anonymiteetin heikkenemisestä seuraa, että ajankohtaisen henkilötietojen käsittelyä koskevan sääntelyn merkitys korostuu, kun yhä useammin niin julkiset kuin yksityiset toimijat käsittelevät toimintansa yhteydessä henkilötietoja.[6] Euroopan unionissa henkilötietojen käsittelyä sääntelee eurooppalainen tietosuojalainsäädäntö[7], jonka pääasiallisina tavoitteina ovat henkilötietojen suoja ja henkilötietojen vapaa liikkuvuus.[8] Tietosuojalainsäädännön tavoitteiden näkökulmasta henkilötiedoiksi tulkittavan tiedon määrän kasvu voi aiheuttaa haasteita. Erittäin suuret tietomassat (*Big data*) mahdollistavat erilaisten kerättyjen tietojen analysoimisen siten, että luonnollisten henkilöiden yksityisyyden ja henkilötietojen suojan toteutumiseen liittyvät riskit kasvavat merkittävästi.[9]

Tästä saattaa jo lähitulevaisuudessa seurata tilanne, jossa jokainen yksittäinen näennäisesti anonyymi tieto voikin olla liitettävissä tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön, jolloin kyseinen tieto päätyy tietosuojalainsäädännön soveltamisalan piiriin.[10] Tällaisessa tilanteessa tulee arvioitavaksi, onko täysin anonyymiä tietoa enää ylipäänsä olemassa. Samalla henkilötietojen anonymisoinnin merkitys korostuu, sillä anonyymi tieto on keskeisessä asemassa muun muassa tieteellisen tutkimuksen ja data-analytiikan näkökulmasta.[11] Tämän kannalta on siis erityisen relevanttia, mikä tulkitaan eurooppalaisessa tietosuojasääntelyssä henkilötiedoksi – ja mikä ei.

Henkilötietoja hyödyntävät toimijat pyrkivät usein säilyttämään henkilöiden anonymiteetin anonymisoimalla kerätyt tiedot siten, ettei yksittäisiä henkilöitä voida tunnistaa tiedoista. Esimerkiksi Transport for London (TfL) alkoi heinäkuussa 2019 kerätä sijaintitietoja Lontoon metron matkustajien älypuhelimista WLAN-verkon välityksellä, jotta tietoja voitaisiin analysoida muun muassa ruuhkien vähentämiseksi ja metroverkon toiminnan optimoimiseksi.[12] TfL:n mukaan kerätyt tiedot anonymisoidaan automaattisesti heti niiden keräämisen jälkeen ja tietojen keräämisen tarkoituksena ei ole selvittää miten yksittäiset ihmiset käyttävät metroverkkoa, vaan ainoastaan kerätä tietoa metroverkon käyttämisestä kokonaisuutena. Metron käyttäjien sijaintitietojen kerääminen perustuu TfL:n

[6] Purtova 2018, s. 78.
[7] Selkeyden vuoksi käytän tässä tutkielmassa termejä ”tietosuoja” ja ”henkilötietojen suoja” synonyymeinä.
[8] Sopimus Euroopan unionin toiminnasta 16 artikla; TSA 1 artikla.
[9] Mayer-Schönberger – Cukier 2013, s. 151–153; Tene – Polonetsky 2013, s. 251–255;
[10] Purtova 2018, s. 77; Ks. myös Bottis – Bouchagiar 2018, s. 195–196. Esimerkiksi terveystietojen ja etenkin genomitietojen käsittelyyn liittyvässä keskustelussa on usein todettu, että täydellistä anonymiteettiä on käytännössä mahdotonta taata. Terveystietojen anonymiteettiin liittyvästä keskustelusta ks. esim. Mostert et al. 2015, s. 958; Sethi – Laurie 2013, s. 187.
[11] Quinn 2017, s. 15–16; Stalla-Bourdillon – Knight 2016, s. 285.
[12] TfL Heinäkuu 2019.

vuonna 2016 toteuttamaan kokeiluun, jossa käytettyä anonymisointitekniikkaa useat tekniset asiantuntijat kuitenkin kritisoivat.[13] Tästä johtuen TfL alkoi käyttämään tietojen tunnistettavuuden poistamiseen kehittyneempää anonymisointiteknologiaa, jonka on tarkoitus taata metron käyttäjien henkilötietojen suoja.

Useiden organisaatioiden hyödyntämä henkilötietojen anonymisointi ei ole kuitenkaan itsestään selvä menetelmä henkilötietojen suojan toteuttamiseksi, sillä anonymisoituihin henkilötietoihin liittyy lähes aina riski anonymisoinnin takaisinmallinnuksesta, jonka seurauksena tietojen kohteena olevat henkilöt voi olla mahdollista tunnistaa. Viime vuosina on esimerkiksi käynyt ilmi, että Twitterin avoimesti julkaistuista metatiedoista[14], yritysten keräämistä verkkosivujen selaustiedoista[15], fitness-sovellusten keräämistä sijaintitiedoista[16] sekä väitetysti anonymisoiduista terveystiedoista[17] on ollut mahdollista tunnistaa yksittäisiä henkilöitä erittäin suurella tarkkuudella.[18] Näin ollen vaikuttaa siltä, että nykyaikaisessa verkottuneessa yhteiskunnassa on yhä vaikeampi pysyä anonyyminä.

1.2 Oikeuslähteet, tutkimuskysymykset ja tutkielman tavoitteet

1.2.1 Oikeuslähteet

Eurooppalaisen tietosuojalainsäädännön perusta on EU-oikeudessa, jossa henkilötietojen suoja tunnustetaan nykyisin itsenäiseksi perusoikeudeksi.[19] Tästä syystä tietosuojasääntely on pohjimmiltaan perusoikeusjuridiikkaa, vaikka tietosuojalainsäädännön tavoitteet käsittävät muitakin näkökulmia kuin luonnollisten henkilöiden henkilötietojen suojan toteutumisen. Tietosuojasääntelyn merkitys on kasvanut Euroopassa huomattavasti viimeisen kahdenkymmenenviiden vuoden aikana, ja sen on jopa sanottu nykyisin olevan yksi EU:n tärkeimmistä arvoista.[20] Tietosuojalainsäädännön tulkintaa ja sys-

[13] Wired 22.5.2019.
[14] Wired 9.7.2018.
[15] Guardian 1.8.2017.
[16] BBC 29.1.2018.
[17] Techscience 29.9.2015.
[18] Guardian 13.7.2018.
[19] Euroopan unionin perusoikeuskirja, 8 artikla. Ks. myös Koillinen 2012, s. 180 ja Lindroos-Hovinheimo 2018, s. 52.
[20] Ks. esim. Petkova 2019, jossa kirjoittaja katsoo yksityisyyden ja henkilötietojen suojan olevan nykyisin Euroopassa samankaltainen ”*leading right*”, kuin mitä sananvapaus on Yhdysvalloissa. Petkovan mukaan sananvapaus ja henkilötietojen suoja ovat oikeuksia, joilla pyritään rakentamaan yhteistä identiteettiä demokraattisessa yhteiskunnassa.

tematisointia vaikeuttaa kuitenkin se, että henkilötietojen suojasta säädellään useassa toisistaan poikkeavassa oikeuslähteessä, joiden keskinäinen hierarkia ei ole täysin selvä.[21] Oikeusjärjestelmän systematiikassa henkilötietojen suojan katsotaan kuuluvan sekä hallinto- että informaatio-oikeuteen, joten se on perinteisessä luokittelussa julkisoikeudellinen oikeudenala.[22] Tästä huolimatta tietosuojalainsäädäntöä sovelletaan yhtä lailla yksityisellä sektorilla tapahtuvaan henkilötietojen käsittelyyn. Näin ollen henkilötietojen suojaa koskevia sääntöjä tulee soveltaa siitä riippumatta, käsitteleekö henkilötietoja julkinen tai yksityinen toimija, kunhan käsittely on ammattimaista.[23]

Henkilötietojen suojan pääasiallisten oikeuslähteiden ollessa peräisin EU-oikeudesta, tutkielman näkökulma on eurooppaoikeudellinen. Tästä johtuen tarkastelen henkilötietojen suojaa eurooppalaisessa kontekstissa pääosin eurooppalaisista oikeuslähteistä käsin. Tärkeimpiä oikeuslähteitä tutkielmassa ovat EU:n sitova primaarinormisto sekä EU:n sekundaarinormit, jotka konkretisoivat niitä periaatteita ja politiikkoja, joita on esitetty primaarinormeissa.[24] Sitovaa primaarinormistoa ovat sopimus Euroopan unionista (SEU), sopimus Euroopan unionin toiminnasta (SEUT) ja Lissabonin sopimuksen myötä primaarinormistoon nostettu Euroopan unionin perusoikeuskirja. Primaarinormistoa konkretisoivia sekundaarinormeja ovat asetukset, direktiivit ja päätökset.[25]

Tärkeinä oikeuslähteinä tutkielman kannalta toimivat lisäksi Euroopan unionin tuomioistuimen (EUT)[26] ratkaisut, sillä nämä ratkaisut ovat etenkin viime vuosina kehittäneet tietosuojalainsäädännön tulkintaa merkittävästi. Unionin tuomioistuimen ratkaisut ovat EU-oikeuden hierarkiassa korkeimmalla, joten niillä on huomattava käytännön merkitys unionin oikeuden soveltamisen kannalta.[27] EUT ratkaisee ennakkoratkaisutuomioistuimena EU-oikeuden tulkinnan kannalta merkittäviä tapauksia, joten sen antamilla ratkaisuilla on merkittäviä vaikutuksia EU-oikeuden yleiseen kehitykseen ja tulkintaan.[28] Näiden ratkaisujen ohella mainittavina oikeuslähteinä tutkielmassa toimivat julkisasia-

[21] Muun muassa Euroopan unionin tuomioistuinten ratkaisut sekä julkisasiamiesten ratkaisuehdotukset asettuvat huonosti perinteiseen kansalliseen oikeuslähteiden hierarkiaan. Tämän tutkielman näkökulmasta näillä oikeuslähteillä on kuitenkin suuri käytännön merkitys. Tarkemmin EU-oikeuden oikeuslähteistä ks. esim. Mäenpää 2011, s. 36–39
[22] Koillinen 2012, s. 172; Korpisaari 2015, s. 997. Eurooppalaisessa oikeuskirjallisuudessa henkilötietojen suoja on usein sijoitettu osaksi hallinto-oikeutta, ks. esim. Hildebrandt 2008, s. 324.
[23] TSA 2(1) artikla. Ks. myös C-25/17 Jehovan todistajat.
[24] Raitio 2016, s. 203.
[25] SEUT 288 artikla; Ojanen 2016, s. 43–44; Raitio 2016, s. 204–205. Asetukset ovat EU:n jäsenvaltioita ja niiden yksityisiä oikeussubjekteja kokonaisuudessaan sitovaa suoraan sovellettavaa oikeutta, jota jäsenvaltioiden ei tarvitse erikseen implementoida. Direktiivit sen sijaan eivät ole suoraan sovellettavaa oikeutta, vaan ne sisältävät jäsenvaltioita sitovan lainsäädäntövelvoitteen, joka jäsenvaltioiden tulee kansallisella lailla täyttää.
[26] Käytän tutkielmassa selkeyden vuoksi myös EUT:ta edeltävästä Euroopan Yhteisön tuomioistuimesta nimitystä EUT.
[27] Sopimus Euroopan unionista 19 artikla.
[28] Sopimus Euroopan unionin toiminnasta 267 artikla. Ks. myös Lock 2015, s. 75–76.

miesten ratkaisuehdotukset, jotka ovat osassa tapauksia EUT:n ratkaisuja seikkaperäisempiä, laajempia ja kattavammin perusteltuja.[29] Julkiasiamiesten ratkaisuehdotukset eivät ole kuitenkaan velvoittavia, vaan ne pikemminkin tarjoavat uusia akateemisia näkökulmia EU-oikeuteen.[30]

Eurooppalaisen henkilötietojen suojaa käsittelevän lainsäädännön tulkinnan kannalta ei voi myöskään sivuuttaa Euroopan tietosuojatyöryhmä WP 29:n[31] lausuntoja ja tulkintakannanottoja. Tietosuojatyöryhmä oli EU:n jäsenvaltioiden tietosuojaviranomaisista koostuva verkosto, jonka henkilötietodirektiivin (95/46/EY)[32] tulkintaa koskevilla lausunnoilla ja kannanotoilla oli ja on edelleen huomattava merkitys eurooppalaisen tietosuojalainsäädännön tulkinnan kannalta. Tietosuojatyöryhmän kannanotot eivät ole velvoittavia oikeuslähteitä, sillä tietosuojatyöryhmästä säännelleen henkilötietodirektiivin 29(1) artiklan mukaan WP 29 oli luonteeltaan neuvoa-antava.[33] Tarkastelen joka tapauksessa tietosuojatyöryhmän lausuntoja ja kannanottoja laajasti tässä tutkielmassa, sillä ne ovat *de facto* vaikuttaneet tietosuojalainsäädännön tulkintaan EU:ssa.[34] Tietosuoja-asetuksen sovellettavaksi tulemisen myötä tietosuojatyöryhmän korvasi Euroopan tietosuojaneuvosto (European Data Protection Board, EDPB), joka on riippumaton EU:n elin.[35] EDPB on ensisijaisesti nähtävissä tietosuojatyöryhmän seuraajana, jolla on neuvoa-antava rooli TSA:n tulkinnassa. Oikeuskirjallisuudessa on kuitenkin arvioitu, että EDPB:n tosiasiallinen asema eurooppalaisen tietosuojalainsäädännön tulkinnassa tulee olemaan huomattavasti tietosuojatyöryhmän neuvoa-antavaa roolia suurempi.[36]

Tulee kuitenkin huomata, että tietosuojatyöryhmän ja myös EDPB:n oikeuslähdeopillinen asema on jossain määrin epäselvä. Tietosuojatyöryhmän lausunnot ja kannanotot eivät ole velvoittavia, mutta niitä tosiasiallisesti noudatetaan useissa organisaatioissa sanantarkasti.[37] Tästä huolimatta EUT ei ole yhdessäkään tietosuojalainsäädännön tulkintaa koskevassa ratkaisussa nimenomaisesti viitannut tie-

[29] Jääskinen 2015, s. 71. Ks. myös SEUT 252 artikla, jonka 2 kohdan mukaan ”*Julkisasiamiesten tehtävänä on täysin puolueettomina ja riippumattomina esittää julkisessa istunnossa perustellut ratkaisuehdotukset asioissa, jotka Euroopan unionin tuomioistuimen perussäännön mukaan vaativat heidän myötävaikutustaan.*”
[30] Ojanen 2016, s. 207.
[31] Henkilötietodirektiiviin mukainen Article 29 working party.
[32] Euroopan parlamentin ja neuvoston direktiivi 95/46/EY annettu 24 päivänä lokakuuta 1995, yksilöiden suojelusta henkilötietojen käsittelyssä ja näiden tietojen vapaasta liikkuvuudesta.
[33] Henkilötietodirektiivi 29(1) artikla.
[34] Zuiderveen Borgesius 2016, s. 259.
[35] TSA 68 artikla.
[36] Poullet 2018, s. 774.
[37] Purtova 2018, s. 43.

tosuojatyöryhmän lausuntoihin, jonka lisäksi unionin tuomioistuin on ratkaissut tapauksia myös tietosuojatyöryhmän kannasta poikkeavalla tavalla.[38] EUT on kuitenkin jossain ratkaisuissa perusteluissaan seurannut tietosuojatyöryhmän asiaa koskevia kannanottoja hyvinkin tarkasti, muttei koskaan tätä nimenomaisesti todennut.[39] Toisaalta julkisasiamiehet ovat usein ratkaisuehdotuksissaan viitanneet tietosuojatyöryhmän lausuntoihin ja hyödyntäneet niitä perusteluidensa tukena.[40]

1.2.2 Tutkimuskysymykset ja tavoitteet

Tutkielmassa tarkastellaan henkilötiedon käsitteen määritelmää eurooppalaisessa tietosuojalainsäädännössä. Henkilötiedon käsitteen tulkinnalla – sillä, millaiset tiedot katsotaan henkilötiedoiksi – on huomattava merkitys niin julkisille kuin yksityisillekin toimijoille Euroopassa. Jos käsittelyn kohteena olevien tietojen katsotaan olevan henkilötietoja, tulee niihin soveltaa tietosuojalainsäädännön velvoitteita. Tarkoituksenani on selvittää, mitä kaikkea henkilötiedon määritelmän alaan eurooppalaisessa tietosuojalainsäädännössä sisältyy. Olennainen osa tätä on sen arvioiminen, mitä kaikkea itse asiassa jää henkilötiedon käsitteen määritelmän ulkopuolelle. Lisäksi kysymys tietosuojalainsäädännön mukaisen henkilötietojen anonymisoinnin mahdollisuudesta, tai toisaalta mahdottomuudesta, liittyy tutkielman aiheeseen olennaisesti. Tästä lähtökohdasta tutkielman keskeiset tutkimuskysymykset muodostuvat kolmesta pääkysymyksestä:

1. Mitä kaikkea sisältyy henkilötiedon käsitteen määritelmään eurooppalaisessa tietosuojalainsäädännössä?

2. Mitä kaikkea jää henkilötiedon käsitteen määritelmän ulkopuolelle?

3. Onko tietosuojalainsäädännön vaatimukset täyttävä henkilötietojen anonymisointi ylipäänsä mahdollista?

Tutkielman tavoitteena on selkeyttää eurooppalaisen tietosuojalainsäädännön mukaisen henkilötiedon käsitteen määritelmää ja sitä, mikä kaikki tieto jää kyseisen määritelmän ulkopuolelle. Lisäksi tavoitteena on tarkastella henkilötietojen anonymisointiin liittyviä haasteita, jotka liittyvät anonyymin tiedon määrittelemisen epävarmuuteen. Tässä yhteydessä arvioin, onko tietosuojalainsäädännön

[38] Purtova 2018, s. 59–60. Ratkaisuja, jossa EUT on ratkaissut tapauksen tietosuojatyöryhmän kannasta poikkeavalla tavalla ks. esim. C-131/12 Google Spain ja C-141/12 YS ym.
[39] Ks. esim. C-434/16 Nowak, kohdat 37–40.
[40] Ks. esim. julkisasiamies Bobekin ratkaisuehdotus asiassa C-40/17 FashionID, kohdat 46 ja 48, sekä julkisasiamies Jääskisen ratkaisuehdotus asiassa C-131/12 Google Spain kohdat 7, 18, 36 ja 48.

vaatimukset täyttävä henkilötietojen anonymisointi ylipäänsä mahdollista. Nämä kysymykset ovat ajankohtaisia, sillä muun muassa erilaiset julkisen sektorin toimijat olisivat halukkaita julkaisemaan tietovarantojaan avoimena datana, jos tämä olisi mahdollista tietosuojalainsäädännön puitteissa.[41] Toisaalta tietojen anonymisointiin pyrkivät myös yksityiset yritykset, joiden intressissä on hyödyntää data-analytiikkaa hallussaan oleviin tietoihin ilman tietosuojalainsäädännön rajoituksia.[42] Näin ollen tutkielman yhtenä tavoitteena on tuottaa tulkintaohjeita rekisterinpitäjinä toimiville tahoille siitä, mitä sekä henkilötiedoilla että anonyymeillä tiedoilla tarkoitetaan, ja miten henkilötietoja voisi olla mahdollista anonymisoida tietosuojalainsäädännön edellytykset täyttäen siten, ettei niiden käsittelystä aiheutuisi tietosuojariskejä rekisteröidyille.

1.3 Metodi

Tutkimuskysymyksiini vastaaminen edellyttää voimassa olevan tietosuojalainsäädännön tulkintaa ja systematisointia, sillä tutkielman pääasiallisena tavoitteena on tämänhetkisen henkilötiedon käsitettä koskevan sääntelyn analysoiminen.[43] Tämän vuoksi tutkielman metodi[44] on oikeusdogmatiikka, jonka sekä käytännöllistä että teoreettista haaraa hyödynnän tutkielmassani: käytännöllinen lainoppi keskittyy oikeusnormien tulkintaan ja teoreettinen lainoppi oikeusnormien systematisointiin.[45] Metodini keskiössä on käytännöllinen lainoppi, jota hyödyntämällä voimassaolevasta oikeudesta on mahdollista muodostaa perusteltuja tulkintakannanottoja.[46] Tutkimukseni kohteena oleva henkilötiedon käsite edellyttää kuitenkin myös käsiteanalyyttista tutkimusotetta, joten hyödynnän käytännöllisen lainopin rinnalla teoreettisen lainopin metodia, jonka systematisointia korostavalla lähestymistavalla on mahdollista analysoida ja jäsentää oikeudellisten käsitteiden sisältöä.[47]

Koska tutkielman näkökulma on EU-oikeudellinen, ei kansallinen lainoppi kuitenkaan yksinään riitä vastaamaan tutkimuskysymyksiini. Tämä johtuu siitä, että monikansallisen *sui generis* -tyyppisen

[41] Ks. esim. Pääministeri Rinteen hallitusohjelma, s. 99, 106 ja 151. Erilaisten tietotyyppien avoimesta julkaisemisesta tarkemmin ks. Zuiderveen Borgesius et al. 2015, s. 2114–2120.
[42] Esimerkiksi Koski et al. 2017, s. 69 raportissa osoitettiin, että informaatio- ja viestintätoimialan yritysten, jotka hyödynsivät avointa dataa ja massadataa innovaatioiden kehittämisessä, liikevaihto kasvoi vuosien 2012–2014 aikana keskimäärin 17 prosenttia enemmän kuin vastaavalla alalla toimivien, mutta dataa hyödyntämättömien yritysten.
[43] Muun muassa Siltala 2003, s. 137, toteaa että ”Oikeustieteen tutkimusmetodi määräytyy tieteenalan, tutkimuskohteen ja valitun tiedonintressin mukaan”.
[44] Metodeilla tarkoitetaan tavanomaisesti tutkimusmenetelmiä, joita hyödyntämällä on mahdollista päätyä tieteellisiin johtopäätöksiin. Ks. esim. Häyhä 1997, s. 24, jossa metodeita kuvataan tutkijan työkaluiksi.
[45] Aarnio 1997, s. 37, Hirvonen 2011, s. 22–25.
[46] Hirvonen 2011, s. 21–22.
[47] Kaisto 2007, s. 17–18. Ks. myös esim. Tuori 2000, s. 309–310. Tuorin mukaan ”Teoreettinen lainoppi systematisoi oikeusjärjestystä kehittelemällä eri oikeudenalojen yleisiä oppeja”.

EU-oikeuden tulkinnassa sääntelyn taustalla vaikuttavilla tavoitteilla on keskeinen merkitys tulkinnan kannalta, ja myös EUT painottaa tulkinnassaan teleologisia argumentteja semioottisten argumenttien sijaan.[48] Näin ollen tutkielman täydentävänä metodina hyödynnän teleologista tulkintaa korostavaa EU-lainoppia. Kansallisen lainopin semanttinen tulkinta on myös yksinään riittämätön lähtökohta EU-oikeuden tulkinnalle, sillä EU:n monikielisyydestä johtuen periaatteessa samoja asioita tarkoittavilla käsitteillä voi olla erilainen merkitys eri kieliversioita sananmukaisesti tulkitsemalla.[49] Lisäksi EU:n hallinnossa muodostunut terminologia voi johtaa siihen, että kansallisessa oikeudessa samat käsitteet tarkoittavat usein eri asiaa kuin EU-kontekstissa.[50]

Tietosuoja-asetuksen suoran sovellettavuuden seurauksena sitä tulee tulkita yhdenmukaisesti kaikissa jäsenvaltioissa, ja tämän vuoksi pelkästä kansallisten metodien hyödyntämisestä tietosuojalainsäädännön tulkinnassa voisi seurata toisistaan poikkeavia ja lainsäädännön harmonisoinnin kannalta ongelmallisia tulkintoja. Näin ollen TSA:n oikeaoppisessa tulkinnassa tulee huomioida EU-oikeuden oikeuslähdeoppi ja EUT:n omaksuma EU-oikeudelle tyypillinen tulkintakäytäntö ja argumentaatio.[51] Tämä tarkoittaa sitä, että tietosuojalainsäädäntöä tulkittaessa tulee tarkastella muun muassa EUT:n soveltuvaa ratkaisukäytäntöä ja tuomioistuimen henkilötietojen suojaa koskevia kannanottoja yksittäistä tapausta laajemmassa kontekstissa.[52] Lisäksi on huomattava, että EU-oikeuden tulkinnassa korostuvat EU:n primaarioikeudessa julkilausutut sääntelyn tavoitteet ja päämäärät, jotka vaikuttavat kaiken EU-lainsäädännön tulkinnan taustalla. Tästä syystä EUT pyrkii tulkitsemaan EU-oikeutta niin sanotun *effet utile* –periaatteen mukaisesti, jolloin tavoitteena on EU-oikeuden mahdollisimman tehokas toteutuminen.[53]

Edellä mainituista syistä tutkielman argumentointi perustuu sekä yleisten EU-oikeuden tavoitteiden että erityisesti tietosuojalainsäädännön tavoitteille. Tällaisen teleologinen tulkinta on tavanomaista

[48] Hirvonen 2011, s. 40. EU-lainsäädännön taustalla olevat poliittiset näkökannat on myös tärkeä huomioida EU-oikeuden tulkinnassa.
[49] Paunio – Lindroos-Hovinheimo 2008, s. 231 ja 234. EUT on käsittelyt EU-oikeuden tulkintaa semioottisten argumenttien kannalta ja todennut muun muassa, että säännösten kieliversioiden vertailu on edellytys EU-oikeuden oikeaoppiselle tulkinnalle. Ks. etenkin ratkaisu C-283/81 CILFIT. Yksityiskohtaisempaa analyysia monikielisestä EU-oikeudesta ks. esim. Paunio – Lindroos-Hovinheimo 2010, s. 410-412.
[50] Paunio 2013, s. 20–21.
[51] EUT hyödyntää ratkaisuissaan muun muassa kielellisiä, systeemisiä, teleologisia ja ylikategorisia argumentteja. Ratkaisukäytännössä kuitenkin useimmiten korostuu teleologiset argumentit. Tarkempaa analyysia EUT:n argumentaatiosta, ks. esim. Bengoetxea 1993, s. 218; Beck 2012, s. 187–188; Conway 2012, s. 131–133.
[52] Sankari 2013, s. 225.
[53] Raitio 2016, s. 51. Effet utile –periaatteesta käytetään myös yleisesti nimitystä tehokkuusperiaate.

haasteellisempaa tietosuojalainsäädännön kontekstissa, sillä jo primaarioikeustasolla on säädetty tietosuojan osalta keskenään jännitteisistä tavoitteista.[54] Tietosuojalainsäädännön tavoitteet – henkilötietojen suoja ja henkilötietojen vapaa liikkuvuus – ovat jännitteisessä suhteessa keskenään, sillä tulkinnassa vastakkain ovat perusoikeudet ja taloudelliset intressit. Tästä johtuen näiden tavoitteiden välisen punninnan huomioiminen on ensisijaisen tärkeää tietosuojalainsäädäntöä tulkittaessa.

Valitun metodin painotukset vaihtelevat hieman tutkimuskysymysten välillä. Ensimmäinen tutkimuskysymys kuuluu suurimmalta osin teoreettisen lainopin alaan, sillä henkilötiedon käsitteen määritelmässä on pitkälti kyse tietosuojalainsäädännön yleisistä opeista. Tämän lisäksi henkilötiedon käsitteen analysoiminen edellyttää semanttista tulkintaa siltä osin, miten käsitteen määritelmä on kielellisesti TSA:ssa ilmaistu. Määritelmän arvioinnissa tulee ensisijaisesti huomioida se mitä määritelmässä sanotaan, mutta toisaalta myös se, mitä siinä ei sanota. Tällä on olennainen merkitys toisen tutkimuskysymykseni kannalta, johon vastaaminen perustuu juuri siihen, miten henkilötiedon määritelmä suhtautuu siihen, mikä jää määritelmän ulkopuolelle. Tähän on mahdollista löytää vastauksia muun muassa arvioimalla tarkkaan, mitä EUT henkilötiedon käsitettä koskevissa tuomiossaan tarkalleen ottaen sanoo, ja millaisen kokonaisuuden ja ratkaisuketjun kyseiset ratkaisut muodostavat.[55]

Kolmas tutkimuskysymykseni on sen sijaan tutkimuskysymyksistäni käytännönläheisin. Pyrkimyksenäni on ensimmäisen ja toisen tutkimuskysymyksen päätelmien perusteella arvioida sitä, millaisissa tilanteissa ja millä edellytyksillä tietosuojalainsäädännön vaatimukset täyttävä henkilötietojen anonymisointi voisi olla mahdollista. Tarkastelen henkilötietojen anonymisointia käytännön esimerkkien kautta ja arvioin anonymisoinnin kautta saavutettavissa olevia lopputuloksia henkilötiedon käsitteen määritelmän ja sen tulkinnan näkökulmasta.

1.4 Tutkielmassa käytetyt käsitteet

Käytän tutkielmassani EU-oikeudellisia käsitteitä samassa merkityksessä kuin nämä käsitteet määritellään EU:n primaarioikeudessa ja henkilötietojen suojaa koskevassa sekundaarilainsäädännössä, eli TSA:ssa.[56] Näin ollen tutkielman avainkäsitteet noudattavat TSA:n 4 artiklan mukaisia määritelmiä.

[54] SEUT 16(1).
[55] Sankari 2013, s. 225. Sankari tarkoittaa ratkaisuketjulla EUT:n samaa aihepiiriä koskevista ratkaisuista muodostuvaa kokonaisuutta, jota vasten yksittäistä vastaavaa asiaa koskevaa ratkaisua tulisi analysoida. Muun muassa tietosuojan kontekstissa EUT:n ratkaisut muodostavat selkeästi ratkaisuketjun, jolloin uusia ratkaisuja tulisi aina tulkita edellisten ratkaisujen valossa.
[56] Henkilötietojen suojasta sähköisen viestinnän alalla säädetään myös sähköisen viestinnän henkilötietodirektiivissä, jonka yksityiskohtaisemman käsittelyn rajaan tietoisesti tutkielman ulkopuolelle. Huomionarvoista on kuitenkin, että EU:ssa on parhaillaan valmisteilla ePrivacy-asetus, joka tulee olemaan *lex specialis* TSA:n suhteen.

Henkilötietojen käsittelyn ja rekisterinpitäjän käsitteet ovat tärkeä määritellä etukäteen, sillä niillä on keskeinen rooli tutkielman tutkimuskysymysten kannalta:

> *Henkilötietojen käsittelyllä* tarkoitetaan TSA:n 4(1)(2) artiklaa vastaavasti mitä tahansa toimintoa tai toimintoja, joita kohdistetaan henkilötietoihin tai henkilötietoja sisältäviin tietojoukkoihin joko automaattista tietojenkäsittelyä käyttäen tai manuaalisesti. Henkilötietojen käsittelyä ovat tietojen kerääminen, tallentaminen, järjestäminen, jäsentäminen, säilyttäminen, muokkaaminen tai muuttaminen, hakeminen, kyseleminen, käyttäminen, tietojen luovuttaminen siirtämällä, levittämällä tai asettamalla ne muutoin saataville, tietojen yhteensovittaminen tai yhdistäminen, rajoittaminen, poistaminen tai tuhoaminen.
>
> *Rekisterinpitäjällä* tarkoitetaan TSA:n 4(1)(7) artiklan mukaisesti luonnollista henkilöä tai oikeushenkilöä, viranomaista, virastoa tai muuta elintä, joka yksin tai yhdessä toisten kanssa määrittelee henkilötietojen käsittelyn tarkoitukset ja keinot.[57] Rekisterinpitäjä on yksi eurooppalaisen tietosuojalainsäädännön kannalta keskeisimmistä käsitteistä, sillä tietosuojavelvoitteiden noudattaminen on rekisterinpitäjän vastuulla.[58]

1.5 Tutkielman rakenne

Tutkielma muodostuu neljästä pääluvusta, joista ensimmäinen on johdanto ja neljäs johtopäätökset. Tutkielman varsinainen tutkimusosa perustuu kahdelle pääluvulle, jotka ovat *Henkilötiedon käsite eurooppalaisessa tietosuojalainsäädännössä* (2. luku) ja *Anonyymit tiedot ja henkilötietojen anonymisointi* (3. luku). Näiden tutkielman keskeisen sisällön muodostavien päälukujen perusteella tutkielmaa on periaatteessa mahdollista tarkastella kaksiosaisena: Tutkielman ensimmäinen osa muodostuu henkilötiedon käsitteen positiivisesta puolesta eli siitä, mikä kaikki tieto sisältyy henkilötiedon käsitteen alaan. Tutkielman toinen osa keskittyy taas henkilötiedon käsitteen kolikon kääntöpuoleen, toisin sanoen siihen, mikä kaikki jää henkilötiedon käsitteen ulkopuolelle.

Tutkielman kahteen päälukuun perustuva rakenne pohjautuu eurooppalaisessa tietosuojalainsäädännössä omaksuttuun systematiikkaan, jossa tiedot ovat joko henkilötietoja, jolloin niihin sovelletaan

[57] Tietosuoja-asetuksen suomenkielisessä versiossa käytetään edelleen hieman harhaanjohtavasti 1980-luvun henkilörekisterilain mukaista käsitteistöä, kuten rekisterinpitäjän käsitettä. On oleellisesta todeta, että TSA:ta sovelletaan henkilötietojen käsittelyyn siitä huolimatta, että muodostuuko niistä rekisteriä. Ks. tähän liittyen esim. Neuvonen 2019, s. 233.
[58] WP 169, s. 2.

tietosuojasääntöjä, tai sitten eivät, jolloin tietosuojasäännöt eivät sovellu. Pääluvut ovat tiiviissä yhteydessä toisiinsa, ne tukevat toisiaan ja niissä viitataan toisiinsa. Tästä syystä periaatteellisesta kaksiosaisuudesta huolimatta tutkielma muodostaa yhtenäisen kokonaisuuden, jonka puitteissa asetettuihin tutkimuskysymyksiin on mahdollista vastata. Tutkielman toisessa luvussa tarkastellaan yksityiskohtaisesti henkilötiedon käsitettä eurooppalaisessa tietosuojalainsäädännössä. Pohjustan aihepiiriä käsittelemällä aluksi tietosuojalainsäädännön tavoitteita, joista etenkin henkilötietojen suojan toteutumisen tavoitteen voidaan katsoa vaikuttaneen merkittävästi henkilötiedon käsitteen tulkintaan.

Toisen luvun pääasiallinen sisältö muodostuu tietosuojatyöryhmän muodostamien henkilötiedon käsitteen neljän osatekijän tarkastelusta, joihin henkilötiedon määritelmä eurooppalaisessa tietosuojalainsäädännössä perustuu. Tarkastelen jokaisen osatekijän yhteydessä kyseisen osatekijän kannalta relevanttia EUT:n ratkaisukäytäntöä, joka on vaikuttanut merkittävästi henkilötiedon käsitteen tulkinnan kautta tietosuojalainsäädännön soveltamisalaan. Lisäksi käsittelen henkilötietojen pseudonymisointia, joka liittyy henkilötietojen suojaa koskevan lainsäädännön systematiikassa henkilötiedon käsitteen näkökulmasta positiiviselle puolelle.

Kolmannessa luvussa tarkastelen anonyymien tietojen konseptia, eli toisin sanoen sitä, mikä jää henkilötiedon käsitteen määritelmän ulkopuolelle. Tässä yhteydessä käsittelen myös henkilötietojen anonymisoinnin konseptia, joka käytännössä perustuu henkilötiedon käsitteen tulkinnalle. Analysoin anonymisointia prosessina, jonka seurauksena henkilötiedoista ei pitäisi olla enää mahdollista tunnistaa luonnollisia henkilöitä ja tarkastelen TSA:n tällaiselle prosessille asettamia vaatimuksia. Käyn myös kursorisesti läpi erilaisia anonymisointitekniikoita, jotka perustuvat erilaisille tilastotieteellisille ja teknisille menetelmille. Tarkastelen myös esimerkkejä puutteellisesti toteutetuista anonymisoinneista, joiden seurauksena tiedot eivät ole tosiasiallisesti olleet anonyymejä tietoja. Neljäs ja samalla viimeinen luku on johtopäätökset, jossa analysoin tutkimuskysymyksiäni toisessa ja kolmannessa luvussa tarkastelemieni oikeuslähteiden valossa.

2 Henkilötiedon käsite eurooppalaisessa tietosuojalainsäädännössä

2.1 Henkilötiedon käsitteen tulkinnan tausta

2.1.1 Henkilötietojen suojan asema EU-oikeudessa

Henkilötietojen suojasta alettiin käydä keskustelua Euroopan unionissa (silloisessa Euroopan yhteisössä) ensimmäistä kertaa 1970-luvun alkupuolella.[59] Varsinainen henkilötietojen suojaa koskeva kehikko alkoi kehittymään kuitenkin 1980-luvulla, kun Taloudellisen kehityksen ja yhteistyön järjestö (OECD) antoi kahdeksan suositusta yksityisyydestä ja henkilötietojen suojasta[60], ja Euroopan yhteisön piirissä laadittiin tietosuojasopimus 108[61], joka on edelleen ainut kansainvälinen sopimus henkilötietojen suojasta.[62] Seuraavien vuosien aikana henkilötietojen suojan merkitys korostui entisestään ja tämän kehityskulun seurauksena hyväksyttiin henkilötietodirektiivi 95/46/EY vuonna 1995.[63] Henkilötietodirektiivi oli merkittävä virstanpylväs eurooppalaisessa tietosuojalainsäädännössä, sillä se oli ensimmäinen yleiseurooppalainen jäsenvaltioita sitova lainsäädäntöinstrumentti henkilötietojen suojasta Euroopassa.

Henkilötietojen suoja kirjattiin EU-oikeuteen itsenäiseksi perusoikeudeksi Euroopan unionin perusoikeuskirjan myötä vuoden 2000 joulukuussa.[64] Perusoikeuskirjan oikeudellinen merkitys unionissa oli kuitenkin pitkään epäselvä, sillä se ei ollut oikeudellisesti jäsenvaltioita sitova, vaan pikemminkin ainoastaan poliittinen julistus perusoikeuksien kunnioittamisesta Euroopan unionissa.[65] Tilanne muuttui vuonna 2009 Lissabonin sopimuksen[66] voimaantulon myötä, kun Euroopan unionin perusoikeuskirja sai saman oikeudellisen arvon kuin EU:n perussopimukset.[67] Tämän seurauksena perusoikeuskirja syrjäyttää EU:n primaarioikeuteen kuuluvana sen kanssa ristiriidassa olevat asetukset ja direktiivit.[68] Euroopan perusoikeuskirja on siitä erityislaatuinen asiakirja, että se on ainut kansainvä-

[59] González Fuster – Gutwirth 2013, s. 535.
[60] OECD 1981.
[61] Yleissopimus yksilöiden suojelusta henkilötietojen automaattisessa tietojenkäsittelyssä (ETS nro 108, SopS 36/1992, tietosuojasopimus).
[62] Kindt 2013, s. 91.
[63] Euroopan parlamentin ja neuvoston direktiivi 95/46/EY annettu 24 päivänä lokakuuta 1995, yksilöiden suojelusta henkilötietojen käsittelyssä ja näiden tietojen vapaasta liikkuvuudesta.
[64] Euroopan unionin perusoikeuskirja 8 artikla; González-Fuster 2014, s. 1–3.
[65] Lynskey 2014, s. 2; González-Fuster 2014, s. 213–214.
[66] Ks. Lissabonin sopimus.
[67] Sopimus Euroopan unionista 6(1) artikla.
[68] Rosas 2011, s. 203–204. Tunnettu esimerkki perusoikeuskirjan asemasta on EUT:n ratkaisu C-293/12 Digital Rights Ireland, jossa unionin tuomioistuin kumosi direktiivin 2006/24/EY perusoikeuskirjan 8 artiklan vastaisena.

linen perus- ja ihmisoikeuksista sääntelevä instrumentti, jossa henkilötietojen suoja erotellaan yksityisyyden suojasta sen sijaan, että sitä pidettäisiin yksityisyyden suojan osana.[69] Tästä syystä perusoikeuskirja herätti jo 2000-luvun alussa laajaa keskustelua henkilötietojen suojasta itsenäisenä perusoikeutena Euroopan unionissa.[70]

Henkilötietojen suojasta säädetään perusoikeuskirjan 8 artiklassa, jonka mukaan:

> *”1. Jokaisella on oikeus henkilötietojensa suojaan.*
>
> *2. Tällaisten tietojen käsittelyn on oltava asianmukaista ja sen on tapahduttava tiettyä tarkoitusta varten ja asianomaisen henkilön suostumuksella tai muun laissa säädetyn oikeuttavan perusteen nojalla. Jokaisella on oikeus tutustua niihin tietoihin, joita hänestä on kerätty, ja saada ne oikaistuksi.*
>
> *3. Riippumaton viranomainen valvoo näiden sääntöjen noudattamista.”*

Perusoikeuskirjan 8 artiklan ja perusoikeuskirjalle Lissabonin sopimuksella annetun primaarioikeudellisen aseman myötä EU-oikeuteen muodostettiin yksityisyyden suojasta erillinen jokaisen oikeus henkilötietojensa suojaan, jolla on ollut huomattava merkitys henkilötietojen suojaa koskevan lainsäädännön kehittymiselle EU:ssa.[71] Muun muassa Lynskey on argumentoinut, että ennen Lissabonin sopimusta EUT ei ottanut ratkaisuissaan kantaa henkilötietodirektiivin 1(1) artiklan henkilötietojen suojaa koskevaan tavoitteeseen, sillä unionilla ei ollut riittävää toimivaltaa perusoikeuksista säätämiseen, vaan tuomioistuin sen sijaan korosti ainoastaan henkilötietodirektiivin asemaa sisämarkkinoiden kehittymisessä.[72] Tämä johtui siitä, että henkilötietodirektiivin oikeudellinen perusta nojasi ainoastaan SEUT 114 artiklaan, eli silloisen EY-sopimuksen 95 artiklaan, eikä EUT:lla ollut oikeudellisia edellytyksiä tulkita henkilötietojen suojaa kuin vain sisämarkkinoiden kehitystä edistävästä näkökulmasta.

---

[69] Euroopan unionin perusoikeuskirja 7 ja 8 artiklat.

[70] Laajempi kysymys henkilötietojen suojasta itsenäisenä perusoikeutena on kuitenkin rajattu tämän tutkielman käsittelyn ulkopuolelle tutkimuskysymysteni painotuksen ja rajallisen sivumäärän vuoksi.

[71] Ks. esim. EUT:n ratkaisut C-92/09 Volker und Markus Schecke ja Eifert ja C-280/08 Deutsche Telekom, joissa tuomioistuin nimenomaisesti käsitteli perusoikeuskirjan 8 artiklan mukaista henkilötietojen suojaa perusoikeutena, jonka nojalla EU:n sekundaarilainsäädäntöä voitiin antaa. Ratkaisussa C-28/08 P Bavarian Lager tuomioistuin taas perusteluissaan erotti henkilötietojen suojan yksityisyyden suojasta ja lausui henkilötietojen suojalla olevan erityinen suojansa. Henkilötietojen suojasta perusoikeutena EU-oikeudessa tarkemmin ks. esim. Bonnici 2014, s. 131–133 ja González-Fuster 2014, s. 247–248

[72] Lynskey 2015, s. 51. Ks. myös Petkova 2019, s. 148–149.

Perusoikeuskirjan ohella EU:n primaarioikeudessa sopimuksessa Euroopan unionin toiminnasta (SEUT) säädetään henkilötietojen suojasta ja sen 16 artiklan mukaan:

> *”1. Jokaisella on oikeus henkilötietojensa suojaan.*
>
> *2. Euroopan parlamentti ja neuvosto antavat tavallista lainsäätämisjärjestystä noudattaen luonnollisten henkilöiden suojaa koskevat säännöt, jotka koskevat unionin toimielinten, elinten ja laitosten sekä jäsenvaltioiden, silloin kun viimeksi mainitut toteuttavat unionin oikeuden soveltamisalaan kuuluvaa toimintaa, suorittamaa henkilötietojen käsittelyä, sekä säännöt, jotka koskevat näiden tietojen vapaata liikkuvuutta. Näiden sääntöjen noudattamista valvoo riippumaton viranomainen.”*

Näin ollen SEUT 16(1) artiklassa säännellään kielellisesti perusoikeuskirjan henkilötietojen suojaa vastaavasta jokaisen oikeudesta henkilötietojensa suojaan. Huomionarvoista on kuitenkin 16 artiklan dualismi: ensimmäisessä kohdassa vahvistetaan ensin henkilötietojen suoja, jonka jälkeen toisessa kohdassa korostetaan henkilötietojen suojan ohella henkilötietojen vapaata liikkuvuutta.[73] Näin ollen EU-oikeudessa primaarioikeustasolla vahvistetaan tietosuojalainsäädännölle kahtalaiset tavoitteet, jotka ovat keskenään jännitteisessä suhteessa.

Henkilötietojen suojan yhteydessä on välttämätöntä käsitellä myös yksityisyyden suojaa. Tämä johtuu siitä, että yksityisyyden suoja on merkittävässä roolissa henkilötietojen suojan näkökulmasta, sillä nämä kaksi perusoikeutta ovat EU-oikeudessa osittain päällekkäisiä.[74] Yksityisyyden suojasta säännellään perusoikeuskirjan 7 artiklassa, jonka mukaan:

> *”Jokaisella on oikeus siihen, että hänen yksityis- ja perhe-elämäänsä, kotiaan ja viestejään kunnioitetaan.”*

Yksityisyyden suoja on perusoikeutena huomattavasti henkilötietojen suojaa vanhempaa perua, ja sitä on aina pidetty merkittävänä ja perustavanlaatuisena oikeutena.[75] Yksityisyys ja yksityisyyden

[73] Hijmans 2016, s. 267–268. EUT:n ratkaisukäytännössä henkilötietojen suoja on kuitenkin varsinkin viime vuosina arvotettu henkilötietojen vapaan liikkuvuuden edelle.
[74] Kokott – Sobotta 2013, s. 222 ja 228.
[75] Ks. esim. Warren – Brandeis 1890, s. 195–197, jossa kirjoittajat luonnehtivat yksityisyyden suojan ydinsisällön olevan ”right to be left alone”.

suoja ovat kuitenkin vaikeasti määriteltäviä ja vahvasti kulttuurisidonnaisia termejä, joita on kenties mahdotonta määritellä tyhjentävästi.[76] Tästä huolimatta yksityisyyden suoja on tärkeä perusoikeus EU-oikeudessa ja sen merkitys on viime vuosina korostunut entisestään, kun yksilöistä kerätään yhä useammin henkilökohtaista tietoa eri yhteyksissä. Merkittävin ero yksityisyyden suoja ja henkilötietojen suojan välillä on siinä, että henkilötietojen suoja ei rajoitu pelkästään henkilön yksityisyyden piirissä oleviin tietoihin, vaan henkilötietojen suojaa koskevia säännöksiä sovelletaan silloinkin, kun käsiteltävät tiedot ovat julkisia.[77]

Perusoikeuskirjan ohella toinen merkittävä yksityisyyden suojaa sääntelevä oikeuslähde Euroopassa on Euroopan ihmisoikeussopimus (EIS).[78] Merkittävä ero perusoikeuskirjan ja EIS:n välillä on, että perusoikeuskirjassa henkilötietojen suoja on yksityisyyden suojasta oma erillinen artiklansa, kun EIS:n sääntelyssä henkilötietojen suoja on osa yksityisyyden suojasta sääntelevää 8 artiklaa.[79] Tästä johtuen henkilötietojen suojan kohde on hieman erilainen perusoikeuskirjan ja EIS:n sääntelyssä, sillä EIS:n 8 artikla suojaa lähtökohtaisesti vain henkilön yksityiselämän piiriin kuuluvalta henkilötietojen käsittelyltä.

Huomionarvoista on kuitenkin, että EIS ei ole EU-oikeutta, vaan se sitoo ihmisoikeussopimuksen sopimusvaltioita siitä riippumatta, onko kyseinen sopimusvaltio Euroopan unionin jäsen, tai mitä EU-oikeus kyseiseen tapaukseen sanoo.[80] Näin ollen EIS:n sääntely ja EU-oikeus ovat toistensa kanssa rinnakkaisia järjestelmiä, jotka tukevat toisiaan perus- ja ihmisoikeuksien turvaamisessa. Edellä mainitut perusoikeuskirjan ja EIS:n säännökset muodostavat perustan henkilötietojen suojalle ja yksityisyydelle Euroopassa. Ne eivät kuitenkaan määrittele keinoja näiden tavoitteiden toteuttamiselle, vaan toimivat ainoastaan korkeimpina oikeuslähteinä, joihin pohjaten yksityiskohtaisempaa lainsäädäntöä henkilötietojen suojasta ja yksityisyydestä on mahdollista laatia.

[76] Solove 2008, s. 1–2.
[77] Korpisaari et al. 2018, s. 5–6.
[78] EIS 8 artiklan 1 kohdan mukaan *”Jokaisella on oikeus nauttia yksityis- ja perhe-elämäänsä, kotiinsa ja kirjevaihtoonsa kohdistuvaa kunnioitusta.”*
[79] Perusoikeuskirja 7 ja 8 artiklat; EIS 8 artikla. EU-oikeuden mukainen henkilötietojen suoja ja EIS:n mukainen yksityisyyden suoja ovat kuitenkin osin päällekkäisiä oikeuksia. Ks. Pellonpää et al. 2018, s. 798–799 ja tarkemmin EIT:n ratkaisut López Ribalda ym. v. Espanja, Antovic ja Mirkovic v. Montenegro ja Vukota-Bojic v. Sveitsi, joissa oli kaikissa kyse luonnollisiin henkilöihin kohdistuvasta videovalvonnasta, jonka yhteydessä henkilöiden oikeuksia suojataan niin EIS:n 8 artiklan kuin perusoikeuskirjan 8 artiklan nojalla.
[80] Ks. kuitenkin SEU 6(2) artikla, jonka mukaan ”*Unioni liittyy ihmisoikeuksien ja perusvapauksien suojaamiseksi tehtyyn eurooppalaiseen yleissopimukseen. Liittyminen ei vaikuta perussopimuksissa määriteltyyn unionin toimivaltaan.*” Näin ollen EU:n primaarioikeudessa EU:lle on periaatteessa määritelty velvoite liittyä EIS:n alaisuuteen.

Ajankohtaisin ja samalla yksityiskohtaisin EU-oikeudellinen säännös koskien henkilötietojen suojaa on 25.5.2016 hyväksytty Euroopan parlamentin ja neuvoston asetus (EU) 2016/679 luonnollisten henkilöiden suojelusta henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta ja direktiivin 95/46/EY kumoamisesta (Yleinen tietosuoja-asetus, TSA)[81], jota alettiin soveltaa kahden vuoden siirtymäajan jälkeen 25.5.2018. TSA on asetuksena suoraan sovellettavaa oikeutta, joka ei pääsääntöisesti[82] jätä jäsenvaltioille harkinnanvaraa asetuksen kansallisessa soveltamisessa ja tulkinnassa. Asetuksen tarkoituksena on yhtenäistää henkilötietojen käsittelyä Euroopassa, sillä henkilötietodirektiivi oli implementoitu jäsenvaltioissa eri tavoin, mikä oli omiaan aiheuttamaan erilaisia käytäntöjä henkilötietojen käsittelylle jäsenvaltioiden välillä.[83] TSA on Euroopan unionin verrattain tehokas ratkaisu sekä henkilötietojen suojan yhtenäistämiseksi että sen toteuttamiseksi jäsenvaltioissa.[84] Asetuksessa määritellään muun muassa henkilötietojen käsittelyä koskevat periaatteet, sekä perusteet joiden nojalla henkilötietoja on mahdollista kerätä.[85] Lisäksi siinä säädetään muun muassa rekisteröityjen oikeuksista heidän henkilötietojensa käsittelyssä.[86]

2.1.2 Tietosuoja-asetuksen tavoitteet

Tietosuoja-asetuksen pääasiallisina tavoitteina ovat sen 1 artiklan mukaan henkilötietojen vapaa liikkuvuus ja henkilötietojen suoja, joten TSA:n tavoitteet vastaavat SEUT 16 artiklassa säädettyjä tavoitteita. Asetuksen 1 artiklan mukaan:

> *”1. Tällä asetuksella vahvistetaan säännöt luonnollisten henkilöiden suojelulle henkilötietojen käsittelyssä sekä säännöt, jotka koskevat henkilötietojen vapaata liikkuvuutta.*
>
> *2. Tällä asetuksella suojellaan luonnollisten henkilöiden perusoikeuksia ja -vapauksia ja erityisesti heidän oikeuttaan henkilötietojen suojaan.*
>
> *3. Henkilötietojen vapaata liikkuvuutta unionin sisällä ei saa rajoittaa eikä kieltää syistä, jotka liittyvät luonnollisten henkilöiden suojeluun henkilötietojen käsittelyssä.”*

[81] Tietosuoja-asetuksesta käytetään usein sen englanninkielistä lyhennettä GDPR (General Data Protection Regulation).
[82] Muun muassa TSA:n 9 artiklassa jäsenvaltioille on kuitenkin jätetty kansallista liikkumavaraa muun muassa sosiaali- ja terveystietojen tietosuojan järjestämisessä.
[83] TSA 9 johdantokappale.
[84] Ks. esim. TSV 2018, s. 6–7, jonka perusteella Suomen tietosuojavaltuutetun toimiston asiamäärät 2,5 kertaistuivat TSA:n myötä ja ihmiset ovat nykyään enemmän tietoisia omista tietosuojaoikeuksistaan.
[85] TSA 5 ja 6 artiklat.
[86] TSA III luku rekisteröityjen oikeudet.

TSA:n ollessa käytännössä merkittävin ja yksityiskohtaisin henkilötietojen suojasta sääntelevä oikeudellinen instrumentti EU:ssa, ovat nämä kaksi tavoitetta samalla eurooppalaisen tietosuojalainsäädännön pääasialliset tavoitteet.[87] Kyseiset tavoitteet ovat siitä haasteellisia, että ne ovat jo lähtökohtaisesti jännitteisessä suhteessa keskenään, sillä henkilötietojen suojalla ja henkilötietojen vapaalla liikkuvuudella pyritään saavuttamaan hyvin toisistaan poikkeavia tavoitteita.[88] Tavoitteet korostavat eurooppalaisessa tietosuojalainsäädännössä omaksuttua tasapainottelua luonnollisten henkilöiden omiin tietoihinsa kohdistuvien oikeuksien ja henkilötietojen liiketoiminnallisen, sisämarkkinoiden kehitystä edistävän hyödyntämisen välillä.[89]

Henkilötietojen suojaa koskeva 1(2) artiklan tavoite on johdettavissa perusoikeuskirjan 8 artiklasta ja SEUT 16(1) artiklasta, joten sen tarkoitus on toteuttaa perusoikeuksien, tässä tapauksessa jokaisen oikeuden henkilötietojensa suojaan, toteutumista.[90] Sen sijaan henkilötietojen vapaata liikkuvuutta koskevalla tavoitteella on tarkoitus edistää tehokkaiden sisämarkkinoiden toimintaa, joiden elinehtona on henkilötietojen liikkuvuus jäsenvaltioiden välillä. Henkilötietojen vapaa liikkuvuus on myös merkittävä osa vuonna 2015 julkaistua Euroopan digitaalisten sisämarkkinoiden strategiaa, joka on taas osa laajempaa Eurooppa 2020 -strategiaa.[91] Näin ollen henkilötietojen vapaata liikkuvuutta koskevan tavoitteen tarkoitus on EU:n taloudellisen toiminnan ja kilpailukyvyn vahvistamisessa.[92]

[87] Vrt. TSA:n tavoitteita henkilötietodirektiivin 1 artiklassa säädettyihin tavoitteisiin: ”*1) Tämän direktiivin mukaisesti jäsenvaltioiden on henkilötietojen käsittelyssä turvattava yksilöille heidän perusoikeutensa ja -vapautensa ja erityisesti heidän oikeutensa yksityisyyteen. 2) Jäsenvaltiot eivät voi rajoittaa tai kieltää henkilötietojen vapaata liikkuvuutta jäsenvaltioiden välillä syistä, jotka liittyvät 1 kohdan mukaisesti turvattavaan suojaan”*. TSA:n tavoitteet ovat pääosin samat kuin henkilötietodirektiivissä asetetut tavoitteet sillä merkittävällä poikkeuksella, että kun henkilötietodirektiivin 1(1) artiklassa korostetaan yksilöiden oikeutta yksityisyyteen, korostetaan TSA:n vastaavassa 1(2) artiklassa luonnollisten henkilöiden oikeutta henkilötietojen suojaan. Tämä heijastaa perusoikeuskirjan mukaista henkilötietojen suojan itsenäistä asemaa yksityisyyden suojaan verrattuna.
[88] Lynskey 2015, s. 46.
[89] Korpisaari et al. 2018, s. 34–35.
[90] Reding 2012, s. 120.
[91] Digital single market; Europe 2020 strategy. Lisäksi huomionarvoista on, että henkilötietoja tarvitaan esimerkiksi teknologisten ratkaisujen, kuten erilaisten tekoälyn sovellusten kehittämiseen. Tästä näkökulmasta henkilötietojen vapaalla liikkuvuudella EU:n jäsenvaltioiden välillä on unionin talouskasvun vahvistamista laajempi merkitys. Sen voidaan katsoa olevan yksi edellytys sille, ettei EU jää tekoälyn kehityksessä muusta maailmasta jälkeen.
[92] Hijmans 2016, s. 267–268 mukaan SEUT 16(2) artiklalla ei kuitenkaan välttämättä ole tosiasiallista vaikutusta tietosuojalainsäädännön tavoitteisiin, vaan 2 kohta olisi pikemminkin jäänne henkilötietodirektiivin 1(2) artiklan tavoitteista ja direktiivin 3 johdantokappaleesta, jossa lausuttiin *”3) sellaisten sisämarkkinoiden toteuttaminen ja toiminta, joilla taataan tavaroiden, henkilöiden, palvelujen ja pääomien vapaa liikkuvuus […] edellyttävät paitsi sitä, että henkilötietoja voidaan vapaasti siirtää jäsenvaltioiden sisällä, myös yksilöiden perusoikeuksien turvaamista”.* Vrt. kuitenkin esim. EUT:n ratkaisu C-362/14 Schrems, jossa tuomioistuin korosti henkilötietodirektiivin henkilötietojen vapaata liikkuvuutta koskevan tavoitteen olemassaoloa.

Tällaista aiemmin henkilötietodirektiivissä ja sittemmin TSA:ssa omaksuttua tavoitteiden dualismia on oikeuskirjallisuudessa kritisoitu sekavaksi ja vaikeasti ymmärrettäväksi.[93] Tämän voidaankin katsoa olevan yksi tietosuoja-asetuksen heikkouksista: tavoitteiden ristiriitaisuus ja niiden välinen jännite voivat vaikeuttaa asetuksen tulkintaa ja sitä kautta oikeusvarmuutta.[94] Eri jäsenvaltioissa TSA:ta voidaan alkaa tulkita eri tavalla siksi, että yhdessä jäsenvaltiossa asetusta tulkitaan henkilötietojen suojaa koskevan tavoitteen näkökulmasta, kun toisessa tulkinnan lähtökohtana pidetään henkilötietojen vapaan liikkuvuuden tavoitteen mukaisia taloudellisia intressejä.[95] Tästä voi seurata erilaisia tietosuojaviranomaisten ratkaisuja jäsenvaltioiden välillä, mikä taas voi kannustaa yrityksiä *forum shopping* –näkökulmien huomioimiseen.[96] Tätä ei voida pitää yleisen EU-oikeuden eikä TSA:n tavoitteiden mukaisena. TSA:n tavoitteet esitetään 1 artiklassa samanarvoisina, mutta EUT:n tietosuojalainsäädännön tulkinta on 2010-luvulla korostanut henkilötietojen suojaa huomattavasti henkilötietojen vapaata liikkuvuutta enemmän.[97] Tämä lienee toisaalta perusteltua sen vuoksi, että henkilötietojen suoja on EU-oikeudessa perusoikeus toisin kuin henkilötietojen vapaa liikkuvuus.

Tietosuoja-asetuksen tavoitteet vaikuttavat sen arvioinnissa, miten TSA:n soveltumisala määräytyy. Moderni tietosuojalainsäädäntö korostaa perusoikeuksien tehokasta toteutumista ja tästä syystä TSA nähdään useimmiten ainoastaan henkilötietojen suojaa edistävänä lainsäädäntönä. Henkilötietojen suojan tehokas toteutuminen edellyttää sitä, että lainsäädäntö tunnistaa tilanteet, joissa luonnollisia henkilöitä tulee suojata heidän tietojensa käsittelyssä. Tästä syystä henkilötiedon käsitteen määritelmä, eli toisin sanoen sen määritteleminen, millaisten tietojen käsittely saa aikaan tietosuojalainsäädännön soveltumisen, on ensiarvoisen tärkeää henkilötietojen suojan näkökulmasta. Näin ollen henkilötiedon käsitettä kussakin tilanteessa arvioitaessa tulisi ottaa huomioon tietojen käsittelyn vaikutukset henkilötietojen suojan kannalta ja pohtia, mitä vaikutuksia tietosuojalainsäädännön soveltamatta jättämisellä voisi olla.[98]

---

[93] Esimerkiksi Lynskeyn 2015, s. 87–88 mukaan tavoitteiden välinen suhde on vielä monimutkaisempi kuin miltä se vaikuttaa.

[94] Huomionarvoista on, että varsinkin EUT on tulkinnut tietosuojalainsäädännön tavoitteita vaihtelevasti. Ratkaisuissa C-139/01 Österreichischer Rundfunk and Others ja C-101/01 Bodil Lindqvist EUT korosti erityisesti sisämarkkinoiden kehitystä henkilötietodirektiivin tulkinnassa. Kuitenkin myöhemmin ratkaisuissa C-73/07 Satakunnan markkinapörssi Oy ja Satamedia Oy ja C-275/06 Promusicae EUT:n ratkaisut mahdollistivat pikemminkin jäsenvaltioiden sisämarkkinoiden kehitystä kansallisella lainsäädännöllä rajoittavan tulkinnan. Viimeaikaisessa ratkaisukäytännössään EUT on taas poikkeuksetta arvottanut henkilötietojen suojan henkilötietojen vapaan liikkuvuuden edelle. Ks. esim. ratkaisut C-131/12 Google Spain, C-362/14 Schrems ja C-210/16 Wirtschaftsakademie.

[95] Lynskey 2015, s. 88.

[96] WP 244 rev 01, s. 8. Toisaalta TSA ei mahdollista forum shopping –menettelyä, sillä jäsenvaltioiden tietosuojaviranomaiset määrittelevät objektiivisen arvioinnin perusteella tietyn yrityksen päätoimipaikan, ja tämän jäsenvaltion viranomaiset toimivat kyseisen yrityksen tietosuojavastuista vastaavina viranomaisina.

[97] Ks. esim. julkisasiamies Bobekin ratkaisuehdotus asiassa C-40/17 FashionID, kohta 72, jossa julkisasiamies Bobek totesi, että viimeaikaista EUT:n käytäntöä on leimannut pyrkimys varmistaa tehokas henkilötietojen suoja.

[98] WP 136, s. 4.

2.2 Henkilötiedon käsitteen määritelmä

2.2.1 Yleistä henkilötiedon käsitteestä

Eurooppalaisessa tietosuojalainsäädännössä henkilötiedon määritelmä on erittäin laaja.[99] Tietosuoja-asetuksen 4 artiklan 1 kohdan 1 alakohdassa on henkilötiedon määritelmä, jonka mukaan henkilötiedolla tarkoitetaan:

> *”kaikkia tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön, jäljempänä ’rekisteröity’, liittyviä tietoja; tunnistettavissa olevana pidetään luonnollista henkilöä, joka voidaan suoraan tai epäsuorasti tunnistaa erityisesti tunnistetietojen, kuten nimen, henkilötunnuksen, sijaintitiedon, verkkotunnistetietojen taikka yhden tai useamman hänelle tunnusomaisen fyysisen, fysiologisen, geneettisen, psyykkisen, taloudellisen, kulttuurillisen tai sosiaalisen tekijän perusteella.”*[100]

Kuten määritelmästä käy ilmi, *kaikki tieto* katsotaan henkilötiedoksi, jos se *liittyy* jollain tapaa *tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön*.[101] Henkilötiedon määritelmä on tarkoituksella säädetty laajaksi, jotta henkilötietojen suojaa koskevan lainsäädännön kiertäminen ei olisi mahdollista ja sääntely kattaisi myös soveltamisalan ”varjoalueet” sekä pystyisi ennakoimaan teknologian kehitystä.[102] Tämä myös alleviivaa koko tietosuojalainsäädännön perimmäistä tarkoitusta: luonnollisten henkilöiden henkilötietojen suojan ja yksityisyyden kunnioittamista henkilötietojen käsittelyssä.[103] Niin henkilötietodirektiivin kuin TSA:n mukainen henkilötiedon määritelmä perustuu

---

[99] Euroopan komissio perusteli laajaa henkilötiedon määritelmää sekä alkuperäisessä että muokatussa ehdotuksessaan henkilötietodirektiivin henkilötiedon käsitettä koskevaksi 2 artiklaksi muun muassa tietosuojasopimuksen laajalla henkilötiedon määritelmällä. Ks. tarkemmin Euroopan komissio 1990, s. 19 ja Euroopan komissio 1992, s. 10.

[100] Ks. vastaavasti asetusta edeltävän henkilötietodirektiivin 2(1)(1) kohdassa henkilötiedon käsitteellä oli määritelty tarkoitettavan *”kaikenlaisia tunnistettua tai tunnistettavissa olevaa luonnollista henkilöä (”rekisteröity”) koskevia tietoja; tunnistettavissa olevana pidetään henkilöä, joka voidaan suoraan tai epäsuorasti tunnistaa, erityisesti henkilönumeron taikka yhden tai useamman hänelle tunnusomaisen fyysisen, fysiologisen, psyykkisen, taloudellisen, kulttuurillisen tai sosiaalisen tekijän perusteella.”* Näin ollen henkilötiedon käsitteen ydinsisältö pysyi TSA:n määritelmässä samana, jonka lisäksi siihen sisällytettiin tiettyjä erityisiä tunnistetietoja koskevat maininnat, kuten sijaintitiedot ja verkkotunnistetiedot. Lisäksi sekä henkilötietodirektiivin että TSA:n mukaista henkilötiedon määritelmää tulkittaessa tulee huomioida siinä käytetty sana ”erityisesti” ennen erilaisten tunnistetietojen listaamista, sillä se osoittaa, että tunnistelista ei ole tyhjentävä.

[101] EU-oikeudessa henkilötiedon käsite on laajempi kuin Yhdysvalloissa käytetty vastine *Personal identifiable information* (PII). Yhdysvalloissa käytettävästä PII-käsitteestä tarkemmin ks. esim. Schwartz – Solove 2011.

[102] WP 136, s. 4–6.

[103] Ks. tietosuojasopimus, jonka 1 artiklan mukaan ”*The purpose of this Convention is to protect every individual, whatever his or her nationality or residence, with regard to the processing of their personal data, thereby contributing to respect for his or her human rights and fundamental freedoms, and in particular the right to privacy.*”. Vrt. kuitenkin tietosuojasopimuksen 1 artiklan mukaista yksitavoitteisuutta SEUT 16(1) artiklan ja TSA 1 artiklan dualistisiin tavoitteisiin. Näin ollen on aiheellista kysyä, voiko eurooppalaisella tietosuojasääntelyllä olla enää vain yhtä perimmäistä tarkoitusta.

edelleen tietosuojasopimuksen mukaiselle henkilötiedon määritelmälle[104], jonka laatimisen taustalla vaikutti pääasiallisesti pyrkimys henkilötietojen suojan tehokkaaseen toteutumiseen.[105] Näin ollen henkilötiedon määritelmän tulkinnassa vaikuttaisi edelleen ensisijaisesti korostuvan henkilötietojen suoja.

Henkilötiedon määritelmä pysyi käytännössä samana, kun henkilötietodirektiivin piiristä siirryttiin tietosuoja-asetuksen sääntelyn piiriin. Ainoana erona henkilötietodirektiivin määritelmään verrattuna on, että TSA:n mukaisessa määritelmässä mainitaan eksplisiittisesti eräitä uuden teknologian mahdollistamia tunnistamiskeinoja.[106] Käsitteen määritelmän pysymistä samana puoltaa lisäksi se, että Euroopan unionin julkisasiamies Kokott lausui 20. heinäkuuta 2017 antamassaan ratkaisuehdotuksessa tapaukseen C-434/16 Peter Nowak v Data Protection Commissioner[107], ”*Although the Data Protection Directive will shortly be repealed by the General Data Protection Regulation, which is not yet applicable, the latter will not affect the concept of personal data*”.[108] Näin ollen voidaan todeta, että henkilötiedon käsitteen tulkinta ei käytännössä muuttunut TSA:n myötä.[109]

Koska henkilötiedon käsite pysyi lähes samansisältöisenä TSA:n voimaantulon myötä, on Euroopan tietosuojatyöryhmä WP 29:n lausunto 4/2007 henkilötietojen käsitteestä (WP 136) edelleen relevantti henkilötiedon käsitteen määrittelyssä, vaikka se on laadittu asetusta edeltävää henkilötietodirektiiviä silmällä pitäen.[110] Lausunnossa henkilötiedon käsite pilkotaan neljään toistensa varaan rakentuneeseen ja toisiinsa vahvasti kietoutuneeseen osatekijään, jotka ovat 1) *kaikenlaiset tiedot*, 2) *koskeva*,

[104] Ks. Euroopan komissio 1990, s. 19, jossa komissio toteaa, että "Personal data". As in Convention 108, a broad definition is adopted in order to cover all information which may be linked to an individual. Depending on the use to which it is put, any item of data relating to an individual, harmless though it may seem, may be sensitive (e.g. a mere postal address). In order to avoid a situation in which means of indirect identification make it possible to circumvent this definition, it is stated that an identifiable individual is an individual who can be identified by reference to a number or a similar identifying particular”.

[105] Euroopan komissio 1992, s. 9. Ks. myös tietosuojasopimuksen 2a artikla, jonka mukaisen henkilötiedon määritelmän mukaan *"personal data" means any information relating to an identified or identifiable individual ("data subject").”* TSA:n mukainen henkilötiedon määritelmä on ydinsisällöltään lähes identtinen tietosuojasopimuksen henkilötiedon määritelmän kanssa.

[106] Korpisaari et al. 2018, s. 49–50.

[107] Käsittelen ratkaisua C-434/16 yksityiskohtaisesti kappaleessa 2.2.3.

[108] Julkisasiamies Kokottin ratkaisuehdotus asiassa C-434/16 Nowak, kohta 3.

[109] Toisin sanoen henkilötietodirektiivin mukainen henkilötiedon määritelmä osoittautui hyvin aikaa kestäväksi, kun eurooppalainen lainsäätäjä ei päätynyt sitä muuttamaan kahdenkymmenen vuoden jälkeenkään. Ks. tähän liittyen esim. de Hert – Papakonstantinou 2016, s. 183.

[110] Korpisaari et al. 2018, s. 52. Tämä vuonna 2007 julkaistu lausunto on ohjannut niin kansallisia lainsäätäjiä kuin jäsenvaltioiden tietosuojaviranomaisia tähän päivään asti, joten sen merkitys henkilötiedon käsitteen tulkinnassa on kiistaton. Tietosuojatyöryhmän lausuntojen voidaan katsoa muutenkin vaikuttaneen merkittävästi henkilötietojen suojan soveltamisalan tulkintaan EU:ssa, ja muun muassa julkisasiamiehet ovat usein viitanneet niihin ratkaisuehdotuksissaan. Ks. esim. julkisasiamies Sánchez-bordonan ratkaisuehdotus asiassa C-482/14 Breyer, kohta 16 ja julkisasiamies Sharpstonin ratkaisuehdotus asiassa C-141/12 YS ym., kohta 40.

3) *tunnistettu tai tunnistettavissa oleva* ja 4) *luonnollinen henkilö.*[111] Seuraavaksi tarkastelen henkilötiedon käsitettä seikkaperäisesti näiden neljän osatekijän valossa ja käytän arvioinnissani tietosuojatyöryhmän muodostamaa osatekijöiden jaottelua.

2.2.2 Ensimmäinen osatekijä: Kaikki tiedot

Henkilötiedon käsitteen ensimmäinen osatekijä on ”*kaikki tiedot*”.[112] Tämän osatekijän yhteydessä tietosuojatyöryhmän lausunnossa korostetaan heti alkuun, että lainsäätäjän käyttämä sananmuoto edellyttää laajaa tulkintaa.[113] Tietosuojatyöryhmän mukaan ensimmäistä osatekijää arvioitaessa tulee ottaa huomioon 1) *tiedon luonne,* 2) *tiedon sisältö* ja 3) *tiedon esitystapa.*[114]

Tiedon luonteeseen kuuluvat kaikki objektiiviset ja subjektiiviset tiedot henkilöstä.[115] Objektiivista tietoa ovat tiedot, jotka ovat totta henkilön mielipiteistä tai arvioista riippumatta, kun taas subjektiivista tietoa ovat nimenomaisesti subjektiiviset mielipiteet ja arviot tietystä henkilöstä.[116] Näin ollen objektiivista tietoa on muun muassa tieto siitä, että Matin käsi on murtunut, ja subjektiivista tietoa se, että Tepolla on hyvä huumorintaju. Subjektiivisen tiedon sisällyttäminen henkilötiedon määritelmän alaan on perusteltua siitä syystä, että monenlaisessa henkilön arviointiin perustuvassa toiminnassa käsitellään tällaista tietoa, kun esimerkiksi arvioidaan henkilön soveltuvuutta tiettyyn tehtävään. Tällaisen tiedon rajaamisella henkilötiedon määritelmän ulkopuolelle voisi olla huomattavia negatiivisia seurauksia henkilötietojen suojalle, sillä subjektiivisen tiedon käsittelyllä voi olla merkittäviä vaikutuksia henkilön oikeuksien toteutumisen näkökulmasta.[117] Tietosuojatyöryhmä korostaa subjektiivisten tietojen yhteydessä, että kyseisten tietojen ei tarvitse olla tosia ja niitä tulee käsitellä henkilötietoina niiden mahdollisesta virheellisyydestä huolimatta.[118] Merkittävin huomio tiedon luonteeseen

---

[111] WP 136, s. 6.
[112] Henkilötietodirektiivissä henkilötiedon käsitteen ensimmäinen osatekijä ilmaistiin sanoilla ”kaikenlaiset tiedot”, kun TSA:n suomennoksessa vastaava osatekijä ilmaistaan sanoilla ”kaikki tiedot”. Tämä kohta henkilötiedon käsitteestä muuttui henkilötietodirektiivistä tietosuoja-asetuksen suomennokseen, mutta muutos on mitä ilmeisimmin tehty vain kielellistä sujuvoittamista ajatellen, sillä esimerkiksi englanniksi vastaava kohta oli ”*any information*” niin direktiivissä kuin asetuksessakin. Vastaavasti ranskaksi termi ”*toute information*” pysyi samana siirryttäessä direktiivistä asetuksenmukaiseen määritelmään.
[113] Vrt. alaviitteessä 105 olevaan tietosuojasopimuksen mukaiseen laajaan henkilötiedon määritelmään. Ks. myös Euroopan komissio 1992, s. 10, jossa todetaan, että henkilötiedon käsitteen määritelmän tulisi olla niin yleinen kuin vain on mahdollista, jotta sen alaan sisältyisi kaikki tiedot liittyen tunnistettavissa olevaan henkilöön.
[114] WP 136, s. 6–7.
[115] Ibid, s. 8.
[116] Mittelstadt – Wachter 2019, s. 518.
[117] Jos työhaastattelija voisi esimerkiksi tietosuojalainsäädännön estämättä kertoa muille saman alan työnantajille, että hänen mielestään hakija X oli hajamielinen, voisi tällä olla merkittäviä negatiivisia vaikutuksia X:n työllistymisen kannalta. Tilanteessa sillä ei ole merkitystä, että tämä tieto olisi vain haastattelijan subjektiivinen näkemys.
[118] WP 136, s. 6. Tähän liittyen ks. TSA 15 ja 16 artiklat rekisteröidyn oikeuksista saada pääsy häntä koskeviin tietoihin ja tietojen oikaisemiseen ja poistamiseen.

liittyen on, että tiedon luonteella ei ole merkitystä, kun arvioidaan tiedon sisältymistä henkilötiedon määritelmään. Näin ollen kaikenlaiset tiedot voivat olla henkilötietoja.[119]

Myös EUT on ratkaisussa C-434/16 Nowak ottanut lyhyesti kantaa siihen, mitä henkilötiedon käsitteen ensimmäisellä osatekijällä tarkoitetaan.[120] Ratkaisussa EUT painotti sitä, että henkilötietodirektiivissä säädetty henkilötiedon käsitteen määritelmä on todella laaja ja käsitteessä käytetty ilmaisu ”*kaikki tiedot*” korostaa tätä.[121] Käsitteen alaan sisältyvät kaikki tiedot siitä huolimatta, ovatko ne arkaluonteisia tai yksityisiä, jonka lisäksi tiedot voivat olla objektiivisia tai subjektiivisia, eli myös mielipiteitä tai arvioita, kunhan nämä tiedot liittyvät rekisteröityyn.

Tiedon sisältö on samalla tavalla vähämerkityksellinen seikka tiedon henkilötiedon määritelmän alaan sisältymisen näkökulmasta, kuin tiedon luonne. Tietosuojatyöryhmän lausunnon mukaan kaikki tieto, jota käytetään minkä tahansa tiedon ilmaisemiseen, voi olla henkilötietoa.[122] Edellytyksenä on kuitenkin, että tiedon voidaan katsoa *liittyvän* tunnistettavissa olevaan luonnolliseen henkilöön. Tämä tarkoittaa käytännössä sitä, että henkilötietoja ovat niin henkilön yksityis- ja perhe-elämää[123] koskevat tiedot kuin kaikkea muutakin hänen toimintaansa koskevat tiedot.[124] Tässä yhteydessä myöskään sillä ei ole merkitystä, missä asemassa henkilö kussakin tilanteessa toimii. Tietosuojatyöryhmä korostaa lausunnossaan, että henkilötietojen suoja ei rajoitu pelkästään yksityis- ja perhe-elämän käsitteeseen, vaikka tämä käsite onkin hyvin laaja, sillä Euroopan perusoikeuskirjassa henkilötietojen suoja on erotettu yksityis- ja perhe-elämän kunnioittamista koskevasta 7 artiklasta erilliseksi omaksi 8 artiklakseen.[125] Näin ollen tietosuojatyöryhmä painottaa henkilötietojen suojan merkitystä itsenäisenä perusoikeutena ja sen tulkitsemista yksityisyyden suojasta eroavista lähtökohdista käsin.[126]

[119] Purtova 2018, s. 48.
[120] C-434/16 Nowak.
[121] C-434/16 Nowak, kohdat 33–34. Tämä on toistaiseksi ainut ratkaisu, jossa EUT otti kantaa henkilötiedon käsitteen ensimmäiseen osatekijään, eli siihen, mitä *kaikilla tiedoilla* tarkoitetaan henkilötiedon käsitteen kontekstissa. Henkilötietodirektiivin soveltamisalan laajuuteen liittyen tuomioistuin viittaa ratkaisun C-553/07 Rijkeboer, kohtaan 59, jossa lausutaan erikseen, että direktiivin soveltamisala on erittäin laaja ja henkilötiedoilla tarkoitetaan monenlaisia tietoja.
[122] WP 136, s. 6.
[123] EIS 8 artikla; perusoikeuskirja 7 artikla.
[124] Kuriositeettina Suomen henkilötietodirektiiviä edeltävässä henkilörekisterilaissa henkilöä koskevat julkiset tiedot oli rajattu pois henkilötiedon määritelmästä. Tarkemmin Suomen kansallisessa tietosuojalainsäädännössä siirtymisestä henkilörekisterilaista henkilötietodirektiivin mukaiseen henkilötietolakiin ks. Korhonen 2003.
[125] Vrt. kuitenkin perusoikeuskirjassa omaksuttua ja tietosuojatyöryhmän korostamaa näkemystä henkilötietojen suojasta esimerkiksi Poullet 2018, s. 777–778, jossa Poullet kritisoi EU:n ratkaisua erottaa henkilötietojen suoja yksityisyyden suojasta, sillä Poulletin mukaan henkilötietojen suoja on vain väline yksityisyyden suojan toteuttamiseen. Toisaalta oikeuskirjallisuudessa on myös kannatettu henkilötietojen suojaa itsenäisenä perusoikeutena, ks. esim. Urgessa 2016, s. 531.
[126] WP 136, s. 7.

Tietojen esitystavalla ei ole myöskään ratkaisevaa merkitystä tietojen henkilötiedon määritelmään sisältymistä arvioitaessa. Tiedot voivat olla henkilötietoja riippumatta siitä, millaisessa muodossa ne ovat.[127] Käytännössä tiedot voivat olla esimerkiksi tallennettuina tietokoneelle binäärimuodossa, valokuvina kansioon tai olla kirjoitettuina käsin paperille.[128] Esimerkiksi valokuvat sisältävät henkilötietoja, jos niiden avulla on mahdollista tunnistaa joko kuvat ottanut tai kuvissa esiintyvä henkilö. Tietosuojatyöryhmä käyttää esimerkkinä lapsen piirustusta, jonka lapsi oli piirtänyt perheestään neuropsykiatrisessa tutkimuksessa.[129] Kyseisen piirustuksen voitiin katsoa sisältävän henkilötietoja sekä lapsesta itsestään että hänen vanhemmistaan, sillä piirustus välitti monenlaisia tietoja muun muassa lapsen mielentilasta ja hänen tunteistaan eri perheenjäseniänsä kohtaan. Piirustus sisälsi muun muassa tietoa lapsen psyykkisestä terveydentilasta sekä hänen vanhempiensa käyttäytymisestä, joiden voitiin katsoa olevan henkilötietoja.[130]

Tietojen esitystavan yhteydessä tietosuojatyöryhmä nostaa erikseen esiin biometriset tiedot henkilötiedon käsitteen näkökulmasta.[131] TSA:n 4(1)(14) mukaan biometrisillä tiedoilla tarkoitetaan:

> *”kaikkia luonnollisen henkilön fyysisiin ja fysiologisiin ominaisuuksiin tai käyttäytymiseen liittyvällä teknisellä käsittelyllä saatuja henkilötietoja, kuten kasvokuvia tai sormenjälkitietoja, joiden perusteella kyseinen luonnollinen henkilö voidaan tunnistaa tai kyseisen henkilön tunnistaminen voidaan varmistaa,”.*

Tietosuojatyöryhmän mukaan tavanomaisia biometrisiä tietoja ovat muun muassa sormenjäljet, verkkokalvon rakenne sekä kasvonpiirteet.[132] Biometriset tiedot eroavat muista tavanomaisista henkilö-

[127] Tietojen muoto vaikuttaa kuitenkin siihen, mikä katsotaan TSA 4(1)(2) artiklan mukaiseksi henkilötietojen käsittelyksi. Ks. esim. C-25/17 Jehovan todistajat, jossa oli kyse käsin kirjoitettujen saarnauslistojen tulkitsemisesta henkilötiedoiksi.
[128] WP 136, s. 7. Periaatteessa tiedon muotoa koskeva rajaus on kuitenkin se, että jos henkilöön liittyvä tieto on ainoastaan tietyn henkilön muistissa, tällainen tieto ei ole henkilötietoa. Ks. tähän liittyen Korpisaari 2018, s. 41.
[129] WP 136, s. 8.
[130] Ibid.
[131] Biometrisistä tiedoista ei ollut minkäänlaista sääntelyä tietosuojasopimuksessa eikä henkilötietodirektiivissä, vaan ne tunnistettiin EU:n tietosuojalainsäädännössä eksplisiittisesti vasta TSA:ssa. Tästä huolimatta biometristen tietojen tietosuojasta oli käyty jo pitkään keskustelua oikeuskirjallisuudessa. Ks. esim. Jasserand 2016; Kindt 2013, s. 94–97; Yue Liu 2008, s. 45–48.
[132] WP 136, s. 8. Biometrisiin tietoihin liittyen ks. erityisesti Korja 2016: Biometrinen tunnistaminen ja henkilötietojen suoja: tutkimus biometristen tunnisteiden lainsäädännöllisestä asemasta. Biometrisen tunnistamisen aloista kasvojentunnistus on saanut viime vuosina suurta huomiota mediassa. Ks. esimerkiksi BBC 13.4.2018, jonka mukaan Kiinassa on käytössä satoja miljoonia kasvojentunnistusteknologiaa ja tekoälyä hyödyntäviä valvontakameroita, joiden avulla huhtikuussa 2018 Kiinan poliisi tunnisti etsintäkuulutetun talousrikollisen 60 000 hengen konserttiyleisön joukosta. Vrt. kui-

tiedoista siinä, että ne toimivat sekä tietoina jostain henkilöstä että tietoina, joiden avulla tietty henkilö on mahdollista tunnistaa. Klassisin esimerkki tästä on sormenjäljet, sillä jos tietty henkilö on koskenut johonkin esineeseen, voidaan tästä esineestä otetuista sormenjäljistä tunnistaa kyseinen henkilö.[133] Henkilön sormi ei siis sellaisenaan ole biometristä tietoa, mutta sormesta otettu sormenjälki on. Vastaavasti ihmisten kudosnäytteet ovat tietosuojatyöryhmän mukaan ainoastaan biometristen tietojen lähteitä, eivätkä siis sellaisenaan biometrisiä tietoja.[134]

Näin ollen biometriset tiedot ovat kaksitahoisia: ne toimivat samalla tiedon *sisältönä*, samalla *yhdyssiteenä* luonnolliseen henkilöön. Biometriset tiedot on erikseen määritelty TSA:ssa, sillä ne ovat tietoina ainutlaatuisia ja siten erittäin vahvasti luonnollisen henkilön tunnistavia. Sormenjälkien ohella toisena tunnettuna esimerkkinä biometrisistä tiedoista ovat DNA-tiedot, joiden avulla tietty henkilö on mahdollista tunnistaa sekä ainutkertaisesti että yksiselitteisesti.[135]

Oikeuskirjallisuudessa tietosuojatyöryhmän tulkintaa biometrisistä tiedoista henkilötietoina on kritisoitu epämääräisyydestä, sillä lausunnon perusteluista ei selkeästi ilmene, miten ihmisten kudosnäytteisiin biometrisinä henkilötietoina tulisi suhtautua. Esimerkiksi Purtova on kritisoinut tietosuojatyöryhmän tekemää eroa ihmisten kudosnäytteiden ja lapsen piirustuksen välillä keinotekoiseksi.[136] Tietosuojatyöryhmä katsoo kudosnäytteiden toimivan vain tiedon välittäjinä, jotka eivät ole itsessään ensimmäisen osatekijän mukaisia tietoja, kun taas lapsen piirustus on tiedon lähde sellaisenaan. Purtova pitää tietosuojatyöryhmän lausuntoa tältä osin epäjohdonmukaisena ja näiden kahden esimerkin erottamista toisistaan keinotekoisena, sillä kummassakin tapauksessa tiedot tulee kuitenkin ”erottaa tiedon lähteestä”, että ne ovat henkilötiedon käsitteen ensimmäisen osatekijän mukaisia tietoja.[137]

tenkin New York Times 14.5.2019, jonka mukaan kasvojentunnistusteknologia herättää pelkoja orwellilaisesta valvontayhteiskunnasta ja toukokuussa 2019 San Francisco ensimmäisenä suurena yhdysvaltalaisena kaupunkina kielsi kasvojentunnistusteknologian käytön viranomaisilta.

[133] Myös esimerkiksi henkilön ääni on biometristä henkilötietoa.

[134] WP 136, s. 9.

[135] Ks. tähän liittyen Sophos 18.10.2018, jonka mukaan DNA-tunnistaminen on nykyisin kehittynyt siihen pisteeseen, että henkilöitä on mahdollista tunnistaa DNA:n perusteella, vaikka kyseisen henkilön oma DNA ei olisi yhdessäkään tietokannassa. Yhdysvalloissa vuonna 2018 rikostutkijat onnistuivat tunnistamaan epäillyn rikospaikalta otettujen DNA-näytteiden perusteella, kun he yhdistivät rikospaikalta löytyneen DNA:n olemassa oleviin yleisissä geenitietokannoissa oleviin tietoihin. Tutkijat löysivät epäillyn kaukaisia sukulaisia, joiden DNA:ssa oli riittävästi yhtäläisyyksiä epäillyn DNA:n kanssa, ja näin he lopulta onnistuivat löytämään itse epäillyn. Tämä esimerkki viittaa käytännössä siihen, että mitä useampi ihminen luovuttaa DNA-tietonsa erilaisiin tietokantoihin, sitä todennäköisemmin kenet tahansa on mahdollista tunnistaa DNA-näytteen perusteella.

[136] Purtova 2018, s. 49–50. Kuitenkin oikeuskirjallisuudessa esitetty kritiikki tietosuojatyöryhmän henkilötiedon käsitettä koskevaa lausuntoa kohtaan kohdistuu enimmäkseen tunnistettavuuden konseptiin, ja muihin osatekijöihin kohdistunut kritiikki on ollut huomattavasti vähäisempää.

[137] Purtova 2018, s. 49–50.

Ihmisestä irrotettujen kudosnäytteiden tulkitsemista henkilötiedoiksi on problematisoitu laajemminkin oikeuskirjallisuudessa ja muun muassa Bygrave on kritisoinut eurooppalaisen tietosuojalainsäädännön epäselvää suhtautumista biologisten näytteiden asemaan henkilötietoina.[138] Tämä liittyy olennaisesti siihen, ettei tietosuojatyöryhmän lausunnossa, eikä tietosuojalainsäädännössä ylipäätäänkään ole määriteltyä käsitettä *tieto*, vaan tietoa tarkastellaan vain sen luonteen, sisällön ja esitystavan kautta.[139] Tälle on esitetty kaksi mahdollista syytä: joko tiedon määritelmää pidetään itsestään selvänä, jolloin sitä ei tarvitse määritellä, tai tiedon määrittelemisen katsotaan olevan erittäin vaikea ellei mahdoton tehtävä, jolloin se on tietoisesti jätetty lainsäädännön ulkopuolelle.[140] Näistä mahdollisista syistä huolimatta Bygraven mukaan tiedon, ja etenkin biologisen materiaalin aseman tietona puutteellinen määrittely voi aiheuttaa muun muassa tietosuojalainsäädännön soveltamisalan perusteetonta laajenemista.[141]

Käsite ”tieto” on myös merkityksellinen TSA:n mukaisen henkilötiedon käsitteen määritelmän suomennoksen kannalta. Henkilötiedon käsite on englanniksi ”*personal data*” ja määritelmän ensimmäinen osatekijä on ilmaistu sanoilla ”*any information*”. Tässä huomionarvoista on, että englanniksi termeillä ”*data*” ja ”*information*” tarkoitetaan eri asioita: termillä ”data” viitataan yleisesti sellaiseen jalostamattomaan tietoon, johon ei liity syvällisempää merkitystä, kun taas termi ”i*nformation*” viittaa pikemminkin johonkin todeksi havaittuun asiaan.[142] Suomen kielessä ei ole kuitenkaan suoraa vastinetta termille ”*data*”, vaan TSA:n suomennoksen mukaisessa henkilötiedon määritelmässä vastaavat kohdat ovat ”henkilötieto” ja ”kaikki tiedot”. Tämä voi johtaa muun muassa siihen, että henkilötiedon käsitteen määritelmää vain suomeksi tulkittaessa ensivaikutelmana saattaa olla, että vain todeksi tiedetyt tiedot ovat henkilötietoja, vaikka kuten edellä mainittiin, myös virheelliset ja paikkansapitämättömät tiedot voivat olla henkilötietoja.

[138] Bygrave 2014a, s. 96–97; Bygrave 2014b, s. 126–129.
[139] Tieto on määritelty tieteenaloittain lukuisilla eri tavoilla. Ks. esim. informaatioteorian isänä pidetyn yhdysvaltalaisen matemaatikon Claude Shannonin artikkeli Shannon 1948, *A Mathematical Theory of Communication*, jonka perusteella ”information is the eliminated uncertainty”, italialaisen filosofin Luciano Floridin artikkeli Floridi 2005 *Is Semantic Information Meaningful data?*, jossa tiedon määritelmä on: ”information is data + meaning” ja Mayer-Schönberger – Cukier 2013, jonka mukaan käsitteen “data” määritelmä on ”a description of something that allows it to be recorded, analysed, and reorganized”. En kuitenkaan tässä tutkielmassa perehdy rajallisen laajuuden ja tutkimuskysymysten rajausten vuoksi yksityiskohtaisesti siihen, mitä tiedolla tarkoitetaan.
[140] Purtova 2018, s. 48.
[141] Bygrave 2014a, s. 96–97. Ks. myös S and Marper v UK, kohta 68, jossa EIT totesi lakonisesti yhdellä lauseella, että ihmisen kudosnäytteet: ”constitute personal data within the meaning of the Data Protection Convention as they relate to identified or identifiable individuals”. Otettaessa huomioon, että TSA:n mukainen henkilötiedon määritelmä perustuu tietosuojasopimuksen määritelmälle, on merkittävää, että toinen Euroopan korkeimmista tuomioistuimista on eksplisiittisesti todennut kudosnäytteiden olevan henkilötietoja.
[142] Floridi 2005, s. 352–353.

2.2.3 Toinen osatekijä: Liittyvä

Toinen henkilötiedon käsitteen osatekijä on ”*liittyvä”.*[143] Tietosuojatyöryhmä korostaa toisen osatekijän merkitystä henkilötiedon käsitteen arvioinnissa, sillä tiedon liittyminen tiettyyn henkilöön määrittää usein sen, onko kyseinen tieto henkilötietoa. Kun arvioidaan sitä, että liittyykö jokin tieto tiettyyn henkilöön, tulee erityisesti huomioida kyseisen tiedon suhteet ja yhteydet henkilöön. Tietosuojatyöryhmän mukaan lähtökohtaisesti silloin, kun tiedot *kertovat* jotain henkilöstä, niiden voidaan katsoa *liittyvän* häneen.[144] Tietojen katsotaan liittyvän tiettyyn henkilöön, jos jokin kolmesta tietosuojatyöryhmän nimeämästä ”tekijästä” liittyy tilanteeseen, jossa kyseisiä tietoja käytetään. Nämä tekijät ovat 1) *sisältö*-, 2) *tarkoitus*-, ja 3) *tulostekijä.* Ne ovat keskenään vaihtoehtoisia ja jos yksikin niistä liittyy tilanteeseen, voidaan tietojen katsoa liittyvän tiettyyn henkilöön.

Sisältötekijä on tietosuojatyöryhmän mukaan periaatteessa yksinkertaisin edellytys sille, että tietojen voidaan katsoa liittyvän tiettyyn henkilöön.[145] Tiedon liittymisellä henkilöön sisältönsä puolesta tarkoitetaan yleistä käsitystä siitä, miten jonkin tiedon sisältö kertoo henkilöstä. Esimerkiksi optikon suorittaman näöntarkastuksen tulokset liittyvät selkeästi arvioitavana olleeseen henkilöön sisältönsä puolesta.[146] Tämä tarkoittaa käytännössä sitä, että sisältötekijän mukaiset tiedot liittyvät henkilöön tietosuojalainsäädännön merkityksessä niiden käyttötarkoituksesta huolimatta. Tiedon sisältötekijän arvioinnissa rekisterinpitäjän tai kolmannen tietojen keräämisen tai käsittelyn tarkoituksella, samoin kuin tiedon vaikutuksella rekisteröityyn ei ole merkitystä.

Tarkoitustekijä tarkoittaa sen sijaan sitä, että tiedot ovat henkilötietoja, kun niitä käytetään tai tullaan todennäköisesti käyttämään tietyn henkilön arvioimiseen, tietynlaiseen kohteluun tai henkilöön vaikuttamiseksi.[147] Tarkoitustekijän arvioinnissa on kyse tietojen tarkoituksen arvioinnista: jos tietoja käytetään sellaiseen tarkoitukseen, joka liittyy tiettyyn henkilöön, voidaan näiden tietojen katsoa liittyvän henkilöön tarkoituksensa perusteella. Esimerkiksi vankilan tiloissa olevasta kulunvalvontajär-

[143] Henkilötietodirektiivin suomennoksessa tämä osatekijä oli ilmaistu sanalla ”koskeva”. Kuitenkin tietosuoja-asetuksen englanninkielisessä käännöksessä tämä osatekijä on sama kuin henkilötietodirektiivissä, eli ”*relating to*”. Toisaalta tietosuoja-asetuksen ranskankielisessä versiossa termi muuttui henkilötietodirektiivin termistä ”*concernant*” asetuksen muotoon ”*se rapportant*”. Tämäkin osaltaan korostaa sitä, että sananmukainen tulkinta on ongelmallista monikielisessä oikeudessa.
[144] WP 136, s. 8.
[145] Ibid, s. 10.
[146] Näöntarkastuksen tulokset liittyvät toisaalta jossain tilanteissa myös itse optikkoon, kun esimerkiksi arvioidaan hänen tehokkuuttaan ja taitojaan.
[147] WP 136, s. 10.

jestelmästä voi saada tietoa tiettyjen ovien avaamisesta tiettyinä kellonaikoina ja tällaisia kulunvalvontatietoja on mahdollista tarkastella suhteessa eri asioihin ja henkilöihin. Päiväsaikaan tietyn osaston oven voi avata kuka tahansa vankilassa työskentelevä henkilö, jolloin kyseisen oven avaamista koskevat tiedot ovat mahdollista eri tarkoituksia varten yhdistää vartijaan, sosiaalityöntekijään tai sairaanhoitajaan. Lisäksi ilta-aikaan vankilan ollessa suljettuna, tietyn oven avaaminen on myös yhdistettävissä kunkin iltavuorossa olevan vartijan valvontatarkoitukseen.

Tulostekijä on kolmas peruste sille, että tietojen voidaan katsoa liittyvän tiettyyn henkilöön. Se tarkoittaa sen arvioimista, vaikuttaako tietojen käyttö todennäköisesti[148] henkilön oikeuksiin tai etuihin.[149] Tulostekijän arvioinnissa tulee ottaa huomioon kaikki kyseiseen tapaukseen vaikuttavat olosuhteet ja olennaista tulostekijän arvioimisen kannalta on, että tietojen käsittelystä yksilölle aiheutuvan vaikutuksen ei tarvitse olla suuri, vaan jo pelkästään sen katsovaan riittävän, jos tietojen käsittelyn seurauksena yksilöä saatetaan kohdella eri tavalla kuin muita. Tietosuojatyöryhmä käyttää esimerkkinä yrityksen takseihin asentamaa satelliittipaikannusjärjestelmää, joka kertoo reaaliajassa kunkin taksin sijainnin.[150] Siitä huolimatta, että sijaintietojen käsittelyn tarkoituksena on vain palvelun parantaminen ja polttoaineen säästäminen, eikä tarkoituksena ole arvioida kuljettajien toimintaa, mahdollistaa kyseinen järjestelmä kuljettajien reaaliaikaisen seuraamisen ja arvioinnin. Tästä syystä tällaisella järjestelmällä voi olla merkittävä vaikutus kuljettajien näkökulmasta, joten tulostekijä on tilanteessa läsnä ja järjestelmässä olevia sijaintitietoja on käsiteltävä henkilötietoina.[151]

Toisen osatekijän soveltumisen alarajana voidaan pitää tulostekijän soveltumista tilanteeseen, jolloin tiedot liittyvät tiettyyn henkilöön. Kuten edellä mainittiin, tietosuojatyöryhmän mukaan yhdenkin tekijän soveltuminen tilanteeseen aiheuttaa lähtökohtaisesti sen, että tiedot liittyvät henkilöön. Näin ollen tietojen voidaan näistä kolmesta tekijästä matalimmalla kynnyksellä tulkita liittyvän tiettyyn henkilöön niiden käsittelyn tuloksen seurauksena.[152] Jos tietojen käsittely vaikuttaa todennäköisesti

[148] Purtova 2018, s. 54 kiinnittää huomiota siihen, että WP 29:n henkilötiedon käsitettä koskevan lausunnon tulostekijää käsittelevässä kohdassa lausutaan vain, että vaikuttaako tietojen käyttö ”todennäköisesti” yksilöiden oikeuksiin ja etuihin. Vrt. tietosuoja-asetuksen 26 johdantokappale, jossa tunnistettavuutta arvioitaessa tulisi ottaa huomioon ”kohtuullisen todennäköiset” keinot. Näin ollen tietosuojatyöryhmän lausunnon perusteella tieto liittyisi henkilöön alhaisemmalla kynnyksellä, kuin hänet olisi tiedoista tunnistettavissa.
[149] WP 136, s. 11.
[150] Ibid.
[151] Mielenkiintoinen kysymys on, tulisiko taksiyhtiön käsitellä sijaintitietoja henkilötietoina, jos järjestelmä antaisi automaattisesti kullekin taksille satunnaisen tunnisteen, jonka avulla taksi voitaisiin vain tehokkuuden lisäämiseksi ohjata lähimmän asiakkaan luokse. Jos järjestelmä vielä tämän lisäksi poistaisi sijaintitiedot taksin löydettyä asiakkaan, voisi tämä poistaa tulostekijän tilanteesta.
[152] Purtova 2018, s. 54.

henkilön oikeuksiin tai etuihin, tulisi tietojen tulkita liittyvän tiettyyn luonnolliseen henkilöön. Lisäksi huomionarvoista on, että kaikki tekijät voivat olla läsnä jossain tilanteessa, jolloin tiedot liittyvät henkilöön sisältönsä, tarkoituksensa ja tuloksensa vuoksi.[153]

Esimerkiksi edellä mainitusta optikon näöntarkastuksesta saadut tiedot liittyvät tiettyyn henkilöön sisältönsä perusteella, sillä tämä vastaa yleiskielen käsitystä tiedon sisällön liittymisestä tiettyyn henkilöön. Lisäksi ne liittyvät henkilöön tarkoituksensa vuoksi, koska niitä käytetään henkilön näön arvioimiseen, ja tarpeeksi hyvä tulos on edellytys ajokortin saamiselle. Kolmanneksi ne liittyvät henkilöön tuloksensa vuoksi, sillä tietojen käsittelyn tuloksena voi olla esimerkiksi päätös siitä, että henkilö ei ole oikeutettu ajamaan autoa. Näin ollen sisältötekijän tulkinnassa tulee arvioida tiedon sisällön liittymistä tiettyyn henkilöön, kun tarkoitus- ja tulostekijän tulkinnassa arviointi kohdistuu tiedon käyttötapaan, joista edeltävässä on huomioitava tietojen käsittelyn tavoitteet ja jälkimmäisessä tietojen käsittelyn vaikutus henkilön etuihin ja oikeuksiin. Lopuksi on vielä huomionarvoista, että usein samassa tilanteessa esimerkiksi sisältö- ja tarkoitustekijä saattavat liittyä Minttuun ja tulostekijä Villeen. Tällöin periaatteessa sama tieto voi olla sekä Minttuun että Villeen liittyvää henkilötietoa.

Tietosuojatyöryhmän ohella myös EUT on tarkastellut henkilötiedon käsitteen toista osatekijää vuonna 2014 annetussa ratkaisussaan C-141/12 YS ym., joka oli samalla ensimmäinen EUT:n ratkaisu, jossa unionin tuomioistuin tarkasteli tietosuojalainsäädännön soveltamisen kannalta elintärkeää henkilötiedon käsitettä yksityiskohtaisesti.[154] Tapauksessa oli kyse siitä, *liittyvätkö* oleskelulupahakemuksen käsittelyn yhteydessä vastaavan virkamiehen laatiman päätösluonnoksen liitteenä olevassa muistiossa olevat tiedot oleskeluluvan hakijaan siten, että kyseiset tiedot olisivat henkilötietoja. Liitteenä olevaan muistioon sisältyi oleskeluluvan hakijaa koskevat tiedot ja tapauksen oikeudellinen analyysi. Merkittävimpänä kysymyksenä ratkaisussa oli, onko muistiossa oleva tapauksen oikeudel-

[153] WP 136, s. 11–12.

[154] C-141/12 YS ym., kohta 33. On mielenkiintoista, että EUT tarkasteli henkilötiedon käsitettä yksityiskohtaisesti vasta vuonna 2014 lähes 20 vuotta henkilötietodirektiivin antamisen jälkeen. Useimmissa ratkaisuissa EUT on tyytynyt vain lakonisesti toteamaan tapauksessa kyseessä olevien tietojen olevan henkilötietoja. Esimerkiksi ratkaisun C-101/01 Lindqvist kohdassa 24 todettiin, että henkilön nimen ja puhelinnumeron yhdistelmä on henkilötietoa, ja ratkaisun C-553/07 Rijkeboer, kohdassa 42 henkilön osoitteen todettiin olevan henkilötietoa. Lisäksi ratkaisun C-342/12 Worten, kohdassa 22 todetaan, että henkilön päivittäinen työaika ja tauot katsottiin henkilötiedoiksi, ratkaisussa C-465/00 Österreichischer Rundfunk ym. kohdassa 64 tiettyjen organisaatioiden maksamat rahat ja näiden rahojen vastaanottajat katsottiin henkilötiedoiksi ja ratkaisun C-73/07 Satakunnan Markkinapörssi Oy ja Satamedia Oy, kohdassa 35 henkilöiden tulot todettiin henkilötiedoiksi. Se, että EUT tarkasteli henkilötiedon käsitettä yksityiskohtaisesti vasta ratkaisussa C-141/12 YS ym. johtuu tosin todennäköisesti siitä, että tietojen käsittelyä koskevat tilanteet pysyivät melko pitkään sen verran yksinkertaisina, että kansallisten tuomioistuinten EUT:lle ohjaamat kysymykset eivät edellyttäneet unionin tuomioistuimelta yksityiskohtaista analyysia henkilötiedon käsitteestä. Ks. tähän liittyen Purtova 2018, s. 61.

linen arviointi oleskeluluvan hakijaan liittyvää henkilötietoa. Arvioitaessa tätä kysymystä tietosuojatyöryhmän näkemyksen perusteella, oleskelulupapäätökseen liittyvä oikeudellinen analyysi liittyisi mitä ilmeisimmin tiettyyn henkilöön, sillä tilanteeseen liittyisi sekä tarkoitus- että tulostekijä.[155] Tämä johtuu siitä, että oikeudellista arviointia voidaan katsoa käytettävän tietyn henkilön arvioimiseen tavalla, joka vaikuttaa hänen asemaansa, jonka lisäksi oikeudellisen arvion käsittely vaikuttaa myös kyseisen henkilön etuihin ja oikeuksiin tuloksensa perusteella.

Tapauksessa EUT piti ilmiselvänä, että muistioon sisältyvät oleskeluluvan hakijaa koskevat tiedot ovat henkilötietoja henkilötietodirektiivin merkityksessä.[156] Kuitenkin hieman yllättäen tuomioistuin päätyi siihen lopputulokseen, että tapauksen oikeudellinen arviointi ei ole henkilötietoa, vaikka se *saattaa* sisältää henkilötietoja.[157] Tässä EUT arvioi oikeudellista arviointia henkilötiedon käsitteen näkökulmasta eri tavalla kuin tietosuojatyöryhmä, sillä oikeudellinen arviointi olisi mitä todennäköisimmin tietosuojatyöryhmän kuvaaman toisen osatekijän piirissä, koska sen käsittelyllä voi olla vaikutuksia rekisteröidyn etujen ja oikeuksien kannalta.[158] Ratkaisun perusteluissa viitataan julkisasiamiehen ratkaisuehdotukseen, jossa argumentoitiin oikeudellisen analyysin olevan vain oikeudellisten normien abstraktia tulkintaa, jolloin tällainen oikeudellinen analyysi ei liity tiettyyn henkilöön, tässä tapauksessa oleskeluluvan hakijaan.[159] EUT päätyi tässä ratkaisussa julkisasiamiehen kanssa samalle kannalle.

Tuomioistuin vaikuttaa ratkaisussa tulkitsevan henkilötiedon käsitteen toista osatekijää sen tulkintaa rajoittavalla tavalla, sillä perustelujen argumentoinnin valossa vaikuttaisi siltä, että sekä tarkoitus- että tulostekijät olisivat irrelevantteja sen arvioimisessa, liittyykö tieto kyseiseen henkilöön.[160] Huomionarvoista on myös, että EUT ei myöskään perusteluissaan viittaa nimenomaisesti tietosuojatyöryhmän näkemyksiin, eikä arvioi tilannetta samoilla sanoilla. Tämän on mahdollista katsoa olevan ongelmallista niin tietosuojatyöryhmän auktoriteetin kuin yleisen oikeusvarmuudenkin näkökul-

[155] WP 136, s. 10.
[156] C-141/12 YS ym., kohta 38.
[157] C-141/12 YS ym., kohta 39.
[158] WP 136, s. 11. Ks. myös Wachter – Mittelstadt 2019, s. 526.
[159] C-141/12 YS ym., kohta 40; julkiasiamies Sharpstonin ratkaisuehdotus asiassa C-141/12 YS ym., kohta 59. Huomionarvoista oli, että tällä kannalla olivat myös tapaukseen mielipiteensä antaneet Alankomaiden, Tšekin ja Ranskan hallitukset.
[160] C-141/12 YS ym., kohdat 44–46. Tämä johtuu siitä, että EUT kieltäytyi hyväksymästä hakijoiden argumenttia siitä, että oikeudellisen arvioinnin tulisi olla henkilötietoa, sillä se vaikuttaa oleskeluluvan saantiin. Ks. kuitenkin Zuiderveen Borgesius 2016, s. 260, jossa todetaan, että EUT vastasi kyseisessä ratkaisussa vain rajattuun erityiskysymykseen ja näin ollen ratkaisu ei tarkoittane, ettei henkilöön tuloksensa puolesta liittyvät tiedot olisi henkilötietoja.

masta, sillä EUT ei tuomiossaan nimenomaisesti perustele, miksi se poikkeaa näin merkittävästi tietosuojatyöryhmän kannasta.[161] Ratkaisu vaikuttaa ennen kaikkea perustelujensa valossa siltä, että julkisasiamiehen ratkaisuehdotuksella on ollut huomattava merkitys sekä ratkaisujen perustelujen että sen lopputulokseen kannalta, sillä tuomioistuin seurasi ratkaisuehdotuksen argumentointia hyvin tarkasti.[162] Tästä huolimatta EUT ei kuitenkaan seurannut perusteluissaan julkisasiamiehen omaksumaa selkeää kahtiajakoa tapauksen faktojen ja analyysin välillä, joka olisi voinut potentiaalisesti rajata henkilötiedon käsitteen toisen osatekijän soveltuvuutta vielä merkittävämmin.[163]

Toinen henkilötiedon käsitteen tulkinnan kannalta merkittävä ratkaisu, jossa unionin tuomioistuin tarkasteli tietojen liittymistä henkilöön, on EUT:n joulukuussa 2017 antama ratkaisu C-434/16 Nowak. Tapauksessa oli kyse siitä, ovatko ammatilliseen kokeeseen osallistuneen hakijan laatimat kirjalliset vastaukset ja tarkastajan kyseisiin vastauksiin mahdollisesti tekemät merkinnät kokeeseen osallistuneeseen *liittyviä* henkilötietoja. Ratkaisu liittyi Nowakin aikaisemmin Irlannin tilintarkastajainstituutille lähettämään henkilötietodirektiivin 12 artiklan[164] mukaiseen tietopyyntöön, jossa hän pyysi saada nähtäväksi kaikkia itseään koskevia tietoja, sillä hän oli reputtanut kyseisen instituutin järjestämän kirjallisen kokeen. EUT päätyi tarkastelemaan tapausta kahden ennakkoratkaisukysymyksen valossa:

> *”1) Voidaanko osallistujan ammatillisen kokeen yhteydessä antamissa vastauksissa/antamiksi vastauksiksi tallennettuja tietoja pitää direktiivissä 95/46– – tarkoitettuina henkilötietoina?*
>
> *2) Jos ensimmäiseen kysymykseen vastataan, että kaikkia tai joitakin tällaisia tietoja voidaan pitää direktiivissä tarkoitettuina henkilötietoina, mitkä tekijät ovat merkityksellisiä määriteltäessä sitä, onko tällainen vastauspaperi tietyssä tapauksessa henkilötieto, ja millainen painoarvo tällaisille tekijöille on annettava?”*[165]

Ratkaisussaan EUT analysoi sille asetettuja kysymyksiä kolmen henkilötiedon käsitteen osatekijän perusteella: 1) kaikki tiedot, 2) liittyen ja 3) tunnistettu tai tunnistettavissa oleva. EUT:n mukaan oli

[161] Purtova 2018, s. 60.
[162] C-141/12 YS ym., kohdat 39–41; Julkiasiamies Sharpstonin ratkaisuehdotus asiassa C-141/12 YS ym., kohdat 59–63.
[163] Julkiasiamies Sharpstonin ratkaisuehdotus asiassa C-141/12 YS ym., kohta 56.
[164] Nykyisin vastaavasta tiedonsaantioikeudesta säädetään TSA:n 12 artiklassa.
[165] C-434/16 Nowak, kohta 26.

kuitenkin selkeää, että käsitteen ensimmäinen ja kolmas osatekijä soveltuivat tapaukseen, joten tuomioistuimen perustelut kohdistuivat pääosin henkilötiedon käsitteen toisen osatekijän soveltumiseen tilanteeseen, joka oli monimutkaisempi kysymys.[166] EUT:n perustelujen mukaan tiedot *liittyvät* tiettyyn rekisteröityyn, kun ne koskevat häntä *sisältönsä*, *tarkoituksensa* tai *vaikutuksensa* vuoksi.[167] Näin ollen on huomionarvoista, että aikaisemmasta YS ym. -ratkaisusta poiketen EUT tosiasiallisesti seuraa perusteluissaan tietosuojatyöryhmän lausuntoa henkilötiedon käsitteestä, vaikkei se tätä erikseen perusteluissaan mainitse. Käytännössä ainoana erona unionin tuomioistuin käyttää kolmannesta tietosuojatyöryhmän toiseen osatekijään liittyvästä ”tekijästä” termiä ”vaikutus”, kun tietosuojatyöryhmä käyttää termiä ”tulos”.[168] Kuitenkin osittain erilaisten termien käytöstä huolimatta sekä EUT että tietosuojatyöryhmä arvioivat kyseistä elementtiä samalla tavalla ja päätyvät samankaltaiseen lopputulokseen.

Unionin tuomioistuin päätyi ratkaisemaan tapauksen siten, että kaikki kolme tietosuojatyöryhmän henkilötietojen käsitteen toisen osatekijän mukaista *tekijää* liittyvät tilanteeseen.[169] Sisältötekijä soveltuu, sillä kokeeseen osallistujan vastaukset kertovat hakijan tiedoista ja kompetenssista sekä tietyissä tilanteissa hänen älyllisistä taidoistaan, ajatusprosesseistaan ja arvostelukyvystään, jonka lisäksi käsinkirjoitetut koevastakset kertovat lisäksi hakijan käsialasta.[170] Tarkoitustekijä taas soveltuu, koska ylipäänsä koevastausten keräämisen tarkoituksena on arvioida kokeeseen osallistujan ammattimaista osaamista ja soveltuvuutta kyseiselle alalle.[171] Tulostekijä myös soveltuu tapaukseen, sillä kokeeseen osallistuneen kirjoittamien koevastausten käsittelyn vaikutuksena, toisin sanoen tuloksena on, että pääseekö hakija kyseisestä kokeesta läpi.[172] Tällä on huomattava merkitys kokeeseen osallistuneelle siitä syystä, että se saattaa määrittää pääseekö hakija haluamaansa ammattiin.[173] Näillä perusteilla myös kokeen tarkastajan kokeeseen tekemät kommentit sisältävät kokeeseen osallistujaan liittyviä tietoa, sillä kyseisten kommenttien sisältö kertoo tarkastajan mielipiteitä ja arvioita kokeeseen osallistujasta, erityisesti osallistujan tiedoista ja pätevyydestä.[174] Tästä syystä myös tarkastajan

[166] C-434/16 Nowak, kohdat 29–34.
[167] C-434/16 Nowak, kohta 35.
[168] Ks. WP 136, s. 6–7 ja 10–11. Vrt. kuitenkin C-434/16 Nowak, kohdat 35, 39, 40, 43 ja 44.
[169] C-434/16 Nowak, kohta 40.
[170] C-434/16 Nowak, kohta 37.
[171] Ibid, kohta 38.
[172] Ibid, kohta 39.
[173] Julkisasiamies Kokottin ratkaisuehdotus asiassa C-434/16, kohta 24. Julkisasiamieskin korostaa ratkaisuehdotuksessaan, että kokeen tarkoituksena on arvioida yksilön suoritusta tietystä syystä toisin kuin esimerkiksi kyselyn, jossa kysytään yksilöön liittymättömiä tietoja. Julkisasiamiehen ratkaisuehdotusta lienee kuitenkin tämän perustelun osalta mahdollista kritisoida, sillä yksittäinen henkilö voi olla mahdollista tunnistaa myös sinänsä anonyymin kyselyn vastauksien perusteella, sillä tällaiset vastaukset voivat kertoa henkilön mielipiteistä, arvoista ja näkemyksistä samoin kuin hänelle tunnusomaisista kulttuurillisista tai sosiaalisista tekijöistä. Ks. TSA 4(1)(1) artikla.
[174] C-434/16 Nowak, kohta 42. Lisäksi kuten edellä mainittu, myös subjektiiviset tiedot voivat olla henkilötietoja.

kommentit liittyvät hakijaan, sillä niiden tarkoituksena on raportoida hakijan suoriutumista kokeesta.[175] Lisäksi EUT totesi tarkastajan kommentteihin liittyen vielä erikseen sen olevan tapauksen arvioinnin kannalta merkityksetöntä, että tarkastajan kommentit ovat myös tarkastajaan liittyviä henkilötietoja.[176] Tuomioistuin ei perusteluissaan arvioinut mainintaa tarkemmin, mikä on kyseisten tietojen asema tarkastajaan liittyvinä henkilötietoina, mutta on kuitenkin jokseenkin selvää, että kyseiset tiedot liittyvät tarkastajaan sisältötekijän perusteella, koska kyseiset tiedot kertovat muun muassa tarkastajan subjektiivisesta ajatusprosessista.

EUT:n tuomiossa tarkastellaan myös, mitä siitä seuraisi, jos kokeeseen osallistujan koevastausten ja tarkastajan kommenttien ei tulkittaisi olevan henkilötietoja.[177] Tämän seurauksena kyseisiin tietoihin ei sovellettaisi tietosuojalainsäädäntöä lainkaan, joten kokeeseen osallistuneen vastauksia voisi muun muassa julkaista ja jakaa kolmansille osapuolille vapaasti ilman tietosuojalainsäädännön asettamia rajoituksia.[178] Näin ollen EUT lopulta ratkaisi tapauksen siten, että ammatilliseen kokeeseen osallistuneen kirjalliset vastaukset ja kokeen tarkastajan niihin tekemät kommentit ovat henkilötietodirektiivin 2 artiklan a alakohdan mukaisia kokeeseen osallistuneeseen liittyviä henkilötietoja.[179]

Ratkaisut C-141/12 YS ym. ja C-434/16 Nowak ovat toistaiseksi ainoat EUT:n ratkaisut, joissa tuomioistuin analysoi yksityiskohtaisesti henkilötiedon käsitteen toista osatekijää. Kyseiset ratkaisut ovat yhdessä tarkasteltuna mielenkiintoisia, koska EUT vaikuttaa tulkitsevan käsitteen toista osatekijää rajatummin ensimmäisessä kuin jälkimmäisessä.[180] Tämä johtuu siitä, että unionin tuomioistuin ei antanut YS ym. -ratkaisussa painoarvoa tarkoitus- ja tulostekijöille arvioidessaan, onko oikeudellinen arviointi oleskeluluvan hakijaan liittyvää henkilötietoa. Kuitenkin kolme vuotta myöhemmin annetussa ratkaisussa Nowak, EUT erikseen painotti tarkoitus- ja tulostekijöiden merkitystä arvioitaessa tietojen liittymistä henkilöön ja näin ollen seurasi tietosuojatyöryhmän tulkintaa, vaikkei tuomioistuin käyttänytkään tietosuojatyöryhmän kanssa täysin identtistä terminologiaa.

Toisaalta voidaan kuitenkin argumentoida, etteivät nämä kaksi ratkaisua ole täysin vertailukelpoisia keskenään henkilötiedon käsitteen toisen osatekijän tulkinnan näkökulmasta. YS ym. -ratkaisussa oli

[175] C-434/16 Nowak, kohta 43. Lisäksi kyseiset kommentit luonnollisesti sisältävät myös kokeen tarkastajaan liittyvää tietoa.
[176] Ibid, kohta 44.
[177] C-434/16 Nowak, kohdat 48–49.
[178] Ks. Julkisasiamies Kokottin ratkaisuehdotus asiassa C-434/16, kohta 26, jossa myös julkisasiamies argumentoi, että hakijalla on oikeus kieltää koevastauksiensa tarkistamiseen ja arviointiin liittymätön käsittely.
[179] C-434/16 Nowak, kohta 62.
[180] EUT myös itse toteaa tämän, ks. C-434/16 Nowak, kohta 56.

periaatteessa kyse abstraktista oikeudellisesta arvioinnista, kun taas Nowak -ratkaisussa oli kyse kokeeseen osallistuneen hakijan kirjallisista koevastauksista, joiden käsittelyn ensisijaisena tarkoituksena oli hakijan arvioiminen. Näin ollen YS ym. -ratkaisun voitaneen katsoa liittyvän vain hyvin rajattuun tilanteeseen, jossa on kyse ainoastaan abstraktista tulkinnasta. Tästä huolimatta, jos kyseisiä ratkaisuja tulkitsisi toisistaan irrallaan, olisi henkilötiedon käsitteen toisen osatekijän soveltumista mahdollista tarkastella kahdella erilaisella tavalla. YS ym. -ratkaisun perusteella tieto liittyisi henkilöön vain silloin, kun se liittyisi häneen sisältönsä puolesta. Tämä olisi merkittävästi rajatumpi tulkinta, kuin mitä tietosuojatyöryhmä on omaksunut henkilötiedon käsitettä koskevassa lausunnossaan.

Sen sijaan ratkaisun Nowak perusteella tieto liittyy henkilöön, kun mikä tahansa kolmesta osatekijästä liittyy tilanteeseen, eli kun tiedot liittyvät henkilöön sisältönsä, tarkoituksensa tai tuloksensa puolesta. Ottaen huomioon, että EUT:n ratkaisu tapaukseen Nowak on näistä ratkaisuista uudempi, jonka lisäksi EUT erikseen huomioi Nowak-ratkaisun perusteluissa ratkaisseensa tapauksen eri tavalla kuin aikaisemman YS ym. –tapauksen[181], on unionin tuomioistuimen henkilötiedon käsitteen toista osatekijää koskeva tulkinta tämänhetkisen ratkaisukäytännön valossa tietosuojatyöryhmän tulkintaa vastaava, jolloin tiedot liittyvät henkilöön jo itsessään tulostekijän liittyessä tietoihin.

2.2.4 Kolmas osatekijä: Tunnistettu tai tunnistettavissa oleva

2.2.4.1 Tunnistamisesta ja tunnistettavuudesta yleisesti

Henkilötiedon käsitteen kolmas osatekijä on ”*tunnistettu tai tunnistettavissa oleva”.*[182] Tämä kolmas osatekijä on tutkimuskysymysteni kannalta relevantein, sillä se, liittyykö jokin tieto tunnistettuun tai tunnistettavissa olevaan henkilöön, määrittää käytännössä sen, onko tieto henkilötietoa vai ei.[183] Jos jokin tieto liittyy henkilöön, joka ei ole tunnistettavissa, tällainen tieto ei ole henkilötietoa.[184] Tämä johtuu siitä, että tiedon käsittely ei vaikuta kenenkään henkilön etuihin tai oikeuksiin, jolloin tilanteeseen ei sovellu henkilötietojen suoja perusoikeutena. Tästä syystä tietosuojalainsäädäntöä ei ole myöskään tarkoituksenmukaista soveltaa sellaisen tiedon käsittelyyn, joka ei liity tunnistettavissa olevaan henkilöön.[185]

[181] C-434/16 Nowak, kohta 56.
[182] Kolmas osatekijä on pysynyt TSA:n suomennoksessa samana kuin se oli henkilötietodirektiivissä.
[183] Esayas 2015, s. 2.
[184] TSA 26 johdantokappale.
[185] Perusoikeuskirja 8 artikla; WP 136, s. 5.

Kolmas osatekijä muodostuu kahdesta keskenään vaihtoehtoisesta vaatimuksesta: joko henkilö on 1) *tunnistettu*, tai hän on 2) *tunnistettavissa*. Tämä jaottelu on kuitenkin käytännössä tarpeeton sen arvioinnissa, kuuluuko tieto kolmannen osatekijän alaan. Jos yksittäinen henkilö on tunnistettu, on hän myös tunnistettavissa oleva, ja henkilö on matalammalla kynnyksellä tunnistettavissa oleva kuin tunnistettu.[186] Näin ollen kolmannen osatekijän tulkinnassa tulisi lähtökohtaisesti arvioida ainoastaan sitä, onko henkilö tunnistettavissa.[187]

Tieto liittyy tunnistettuun luonnolliseen henkilöön, jos hänet voidaan erottaa muista samaan ryhmään kuuluvista henkilöistä.[188] On melko suoraviivaista arvioida, liittyykö tieto tunnistettuun henkilöön, joten tämä harvemmin aiheuttaa sen suurempia tulkintaongelmia. Esimerkiksi sanomalehdessä julkaistut tiedot, joissa kerrotaan tulevan pääministerin entisistä rikostuomioista kuvan kera, liittyvät mitä ilmeisimmin tunnistettuun henkilöön. Huomattavasti monimutkaisempaa on sen sijaan arvioida sitä, milloin henkilö on tunnistettavissa, toisin sanoen häntä ei ole tunnistettu, mutta hänet on *mahdollista tunnistaa.* Tämä ei edellytä sitä, että tunnistaminen olisi jo tapahtunut, vaan pelkkä mahdollisuus henkilön tunnistamisesta johtaa tietosuojalainsäädännön soveltumiseen kyseisten tietojen käsittelyssä.[189]

Tietosuojatyöryhmä on todennut luonnollisen henkilön tunnistettavuuden määrittävän sen, tulkitaanko jonkin tiedon kuuluvan kolmannen osatekijän alaan.[190] Tunnistaminen perustuu tavanomaisesti yksittäisten *tunnisteiden* yhdistämiseen tiettyyn henkilöön. TSA:n 4(1)(1) artiklan mukaisessa henkilötiedon määritelmässä todetaan esimerkkeinä tunnisteista *henkilön nimi, henkilötunnus, sijaintitieto, verkkotunnistetieto[191], sekä hänelle tunnusomaiset fyysiset, fysiologiset, geneettiset, psyykkiset, taloudelliset, kulttuurilliset ja sosiaaliset tekijät.* Henkilötiedon määritelmään sisällytetty lista ei ole kuitenkaan tyhjentävä, jonka vuoksi tunnistettavuutta arvioitaessa tulee ottaa huomioon kaikenlaiset *tunnisteet*, jotka on mahdollista yhdistää tiettyyn henkilöön.

---

[186] Esayas 2015, s. 2.
[187] Myös Hintze 2018, s. 89 toteaa, että käytännön ero tietojen välillä, jossa olevat henkilöt on tunnistettu tai tunnistettavissa on ollut käytännössä merkityksetön, sillä tietojen käsittelyyn sovelletaan samoja vaatimuksia molemmissa tilanteissa.
[188] WP 136, s. 12.
[189] Bolognini – Bistolfi 2017, s. 174.
[190] WP 136, s. 12.
[191] Ks. WP 188, jossa tietosuojatyöryhmä linjasi internet-evästeiden olevan vastaavia tunnisteita. Tämä on vahvistettu myös TSA:n 30 johdantokappaleessa, jonka mukaan ”*Luonnolliset henkilöt voidaan yhdistää heidän käyttämiensä laitteiden, sovellusten, työkalujen ja protokollien verkkotunnistetietoihin, kuten IP-osoitteisiin, evästeisiin tai muihin tunnisteisiin, esimerkiksi radiotaajuustunnisteisiin. Näin käyttäjästä voi jäädä jälkiä, joita voidaan käyttää luonnollisten henkilöiden profilointiin ja tunnistamiseen etenkin, kun niitä yhdistetään yksilöllisiin tunnisteisiin ja muihin palvelimelle toimitettuihin tietoihin.”.*

Tässä yhteydessä on erityisen tärkeää todeta henkilötiedon käsitteen olevan teknologianeutraali.[192] Tunnistettavuuden mahdollistavalla teknologialla ei ole merkitystä, jos henkilö on tosiasiallisesti tunnistettavissa. Näin ollen tunnistaminen voi tapahtua sekä toisen henkilön toimesta ”perinteisesti”, kun hän esimerkiksi tunnistaa valokuvassa olevan henkilön, tai tunnistamisen voi suorittaa tietojärjestelmä, joka vertaa henkilöstä otettua kuvaa tuhansiin tietokannassaan oleviin kuviin ja löytää vastaavan kuvan tunnistavine tietoineen.

Tosiasiallisesti tietty henkilö on useimmiten tunnistettavissa jopa yksittäisten tunnisteiden perusteella.[193] Tällaisia tunnisteita ovat muun muassa henkilön pituus, silmien väri tai hänen vaatetyylinsä. Tunnistettavuus ei ole kuitenkaan rajoittunut vain ulkoisiin seikkoihin, ja henkilö on mahdollista tunnistaa myös esimerkiksi sen perusteella, että hän toimii väitöskirjatutkijana.[194] Henkilötiedon käsitteen kolmanteen osatekijään liittyy olennaisesti henkilötiedon määritelmän kohta, jonka mukaan tunnistettavissa olevana pidetään henkilöä, joka voidaan *suoraan tai epäsuorasti* tunnistaa. Tietosuojatyöryhmän mukaan henkilön tunnistamisella suoraan tarkoitetaan henkilön tunnistamista esimerkiksi hänen nimensä perusteella, sillä nimi on tavallisin henkilön tunnistava tekijä.[195] Henkilön tunnistamisella epäsuorasti tarkoitetaan sen sijaan sitä, että henkilö on tunnistettavissa esimerkiksi puhelinnumeron, auton rekisterinumeron, henkilötunnuksen, passin numeron tai TSA:n 4(1)(1) artiklassa uusina kohtina mainittujen sijaintitietojen tai verkkotunnistetietojen perusteella.[196]

Tietosuojatyöryhmän henkilötiedon käsitettä koskevassa lausunnossa oleva jaottelu tunnistamiseen suorasti ja epäsuorasti vaikuttaa perustuvan ainoastaan komission täydennetylle ehdotukselle henkilötietodirektiivin mukaiseksi henkilötiedon käsitteeksi vuodelta 1992, jonka perusteella henkilö olisi

[192] Tästä syystä sillä ei ole käytännössä merkitystä, että TSA:n henkilötiedon määritelmään lisättiin uusina esimerkkeinä tunnisteista sijaintiedot ja verkkotunnistetiedot. Käsitteen teknologianeutraaliudesta johtuen tällaiset tiedot olisivat henkilötiedon käsitteen piirissä ilman eksplisiittistä mainintaakin niistä.
[193] Huomionarvoista on, että tämä lausutaan eksplisiittisesti myös ehdotuksessa tietosuoja-asetukseksi. Ks. Euroopan komissio 2012, jonka 24 johdantokappaleen mukaan ”Verkkopalvelujen käyttäjät voidaan yhdistää heidän käyttämiensä laitteiden, sovellusten, työkalujen ja protokollien internet-tunnisteisiin, kuten IP-osoitteisiin tai evästeisiin. Näin käyttäjästä voi jäädä jälkiä, joita voidaan käyttää yhdessä ainutlaatuisten tunnisteiden ja muiden palvelimille toimitettujen tietojen avulla käyttäjien profilointiin ja tunnistamiseen. Tästä seuraa, että tunnistenumeroita, sijaintitietoja, internet-tunnisteita tai muita erityistekijöitä ei sinänsä tarvitse pitää henkilötietoina kaikissa tilanteissa”. Ks. tähän liittyen myös Schwartz – Solove 2014, s. 886, jossa kirjoittajat argumentoivat TSA:n silloisen ehdotuksen mahdollistavan tunnistamisen arvioinnin aina kuhunkin tilanteeseen räätälöidyllä analyysillä, jonka avulla voisi periaatteessa valita, onko tilanteessa läsnä henkilötietoja.
[194] WP 136, s. 12.
[195] WP 136, s. 12–13. Tietosuojatyöryhmä siteeraa lausunnossaan komission täydennettyä ehdotusta henkilötietodirektiivin henkilötiedon käsitettä koskevaksi 2 artiklaksi. Ks. tarkemmin Euroopan komissio 1992, s. 9. Lausunnossa huomioidaan kuitenkin se, että henkilön nimi ei ole aina sellaisenaan riittävä tietyn henkilön tunnistamiseksi, sillä tunnistettavuuden arvioinnissa tulee ottaa huomioon myös täysin samannimiset henkilöt.
[196] WP 136, s. 12–13.

periaatteessa tunnistettavissa suoraan vain hänen nimensä perusteella ja epäsuorasti kaikkien muiden tunnisteiden perusteella.[197] Vastaava jaottelu suoraan ja epäsuoraan tunnistamiseen on otettu TSA:n 4(1)(1) artiklaan, mutta sen käytännön merkitys ontuu, jos henkilö voidaan tunnistaa suoraan vain hänen nimensä perusteella. Tämä johtuu ennen kaikkea siitä, että nimet eivät ole samalla tavalla ainutlaatuisia, kuin esimerkiksi henkilötunnukset.[198]

Tästä näkökulmasta voi olla tarkoituksenmukaisempaa katsoa henkilön olevan suoraan tunnistettavissa ainoastaan hänen henkilötunnuksensa perusteella hänen nimensä sijaan.[199] Tietosuojatyöryhmä kuitenkin toteaa, että tietyissä tilanteissa henkilön nimi tulee yhdistää muihin henkilöön yhdistettävissä oleviin tunnisteisiin, kuten hänen syntymäpäiväänsä ja osoitteeseensa. Toisaalta, vaikka henkilö tavanomaisesti tunnistetaan hänen nimensä perusteella, ei tämä ole tietyissä tapauksissa välttämätöntä. Esimerkiksi jos henkilön erottaminen joukosta on mahdollista yhdistelemällä tunnisteita, voidaan hänen olevan tunnistettavissa siitä huolimatta, että hänen nimensä säilyy anonyyminä.[200]

Henkilön tunnistamisen arvioinnin kannalta suurin merkitys on tilanteella ja asiayhteydellä. Tavanomaisen etunimen perusteella ei ole mahdollista tunnistaa yksittäistä henkilöä kaupungin asukkaista, mutta nimen perusteella voi hyvinkin olla mahdollista tunnistaa henkilö työpaikan kollegoiden joukosta.[201] Toissijainen tieto, joka ei sellaisenaan riitä henkilön tunnistamiseen, kuten tieto ”punatukkaisesta naisesta metrossa” voi kuitenkin mahdollistaa henkilön tunnistamisen, jos kyseinen henkilö on ainut punatukkainen ja naispuolinen henkilö kyseisessä metrossa.[202]

Tosiasiallisesti henkilön tunnistaminen epäsuorasti tapahtuu useimmiten yksittäisten tunnisteiden yhdistelyllä. Tämä tarkoittaa sitä, että yksittäinen tiedonpalanen ei välttämättä vielä sellaisenaan riitä henkilön tunnistamiseksi, mutta yhdistämällä tämä tunniste muihin tunnisteisiin on mahdollista löy-

[197] Euroopan komissio 1992, s. 9. Oikeuskirjallisuudessa suorien tunnisteiden joukko on katsottu johdonmukaisesti tietosuojatyöryhmän näkemystä laajemmaksi, ks. esim. Hintze 2018, s. 87, jossa kirjoittaja tulkitsee suoriksi tunnisteiksi nimet, puhelinnumerot ja valtioiden kansalaisilleen antamat henkilökohtaiset tunnukset.
[198] Ks. tähän liittyen OECD 2013, s. 52, jossa henkilönumero (”*a civil registration number*”) on tulkittu sellaiseksi tiedoksi, josta henkilö on suoraan tunnistettavissa. Sen sijaan OECD:n näkemyksen mukaan esimerkiksi kotiosoite on tietoa, josta henkilö on ainoastaan epäsuorasti tunnistettavissa.
[199] WP 136, s. 14.
[200] Zuiderveen Borgesius 2016, s. 262.
[201] WP 136, s. 13.
[202] Vastaavasti tutkielman johdannossa esitellyssä esimerkissä Lontoon metron käyttäjien sijaintitietojen keräämisestä yksittäinen henkilö voi olla mahdollista tunnistaa, jos hän on ainut tiettyyn kellonaikaan tietyllä asemalla liikkuva henkilö.

tää sellainen ainutlaatuinen yhdistelmä tunnisteita, jonka perusteella henkilö voi olla tunnistettavissa.[203] Näin ollen yksittäinen tieto saattaa vaikuttaa siltä, että se ei liity tunnistettavissa olevaan henkilöön.[204] Tällainen tieto voisi esimerkiksi olla, että yksi tiedoissa oleva henkilö on poliitikko. Jos tämä tieto on mahdollista yhdistää tietoihin kyseisen henkilön sukupuolesta ja postinumerosta, tulee henkilön tunnistettavuudesta jo huomattavasti todennäköisempää, sillä mahdollisten henkilöiden piiri pienenee merkittävästi. Tällöin henkilö on jo todennäköisemmin erotettavissa joukosta, jolloin hän on mahdollisesti tietosuojalainsäädännön näkökulmasta tunnistettavissa.

Esimerkiksi laajoja niin sanottuja ryhmätason tietoja yhdistelemällä yksittäinen henkilö saattaa olla tunnistettavissa pelkästään syntymäajan, sukupuolen ja postinumeron perusteella.[205] Toisaalta yksittäiset tunnisteet voivat olla sellaisenaankin riittävän ainutlaatuisia henkilöiden tunnistamiseksi ja esimerkiksi tieto, että henkilö on jonkin puolueen puheenjohtaja, riittää useimmiten kyseisen henkilön tunnistamiseen. Näin ollen tärkeä elementti tunnistettavuuden arvioinnissa on kysymys siitä, onko henkilö mahdollista erottaa joukosta.[206] Tämä tarkoittaa sitä, että tunnistettavuus ei välttämättä edellytä henkilön tarkan identiteetin paljastumista, vaan pikemminkin sellaisen ainutlaatuisen tunnisteiden joukon havaitsemista, joka liittyy luonnolliseen henkilöön.[207]

Edellä tarkasteltujen tunnistettavuutta koskevien seikkojen valossa monimutkaiseksi kysymykseksi muodostuu, onko sellainen tieto henkilötietoa, joka liittyy tiettyyn joukkoon henkilöitä vain tietyllä todennäköisyydellä. Jos esimerkiksi kaikki tietyn postinumeron alueella äänestivät eduskuntavaaleissa samaa puoluetta, on tieto alueen äänestyskäyttäytymisestä alueella asuneisiin henkilöihin liittyvä henkilötieto, sillä on vähällä vaivalla selvitettävissä, keitä henkilöitä kyseinen tieto koskee.[208] Huomattavasti vaikeampaa on sen sijaan arvioida, onko henkilötietoa jo pelkästään tieto siitä, että 40

[203] Muun muassa Ison-Britannian kansallinen tietosuojaviranomainen ICO (Information Commissioner's office, ICO) käyttää tällaisesta tunnistamisesta termiä "*jigsaw identification*". Ks. ICO 2012, s. 24.
[204] Esimerkiksi karttatiedot saattavat nopeasti arvioituna vaikuttaa siltä, etteivät ne liity tunnistettavissa olevaan henkilöön. Kuitenkin karttatietojen tarkemmasta tarkastelusta käy useimmiten ilmi, että varsinkin tarkoista ja yksityiskohtaisista kartoista, on mahdollista löytää tietoja, jotka liittyvät tunnistettavissa oleviin henkilöihin. Esimerkiksi useimmissa digitaalisissa kartoissa näkyvät yksittäiset talot, niiden osoitteet sekä talojen numerot ovat helposti yhdistettävissä tunnistettavissa oleviin henkilöihin yhdistämällä tiedot esimerkiksi joko väestötietojärjestelmässä oleviin julkisiin tietoihin tai yhteystieto- ja mediayhtiö Fonectan internetissä julkaisemiin luonnollisten henkilöiden yhteystietoihin. Ks. tarkemmin karttatiedoista henkilötietoina van Loenen et al. 2016.
[205] Sweeney 2002, s. 2. Esimerkiksi Latanya Sweeney onnistui jo vuonna 2002 yhdistämään de-identifioidut potilastiedot Yhdysvaltojen väestötietojärjestelmän äänestäjälistoihin ja todistamaan, että 87% Yhdysvaltojen väestöstä oli mahdollista tunnistaa, jos vain tiesi henkilön syntymäajan, sukupuolen ja postinumeron.
[206] WP 136, s. 12–13. Ks. myös Zuiderveen Borgesius 2016, s. 260.
[207] WP 136, s. 13; Jasserand 2016, s. 203.
[208] Korpisaari et al. 2018, s. 55.

prosenttia alueella asuvista äänesti kyseisissä vaaleissa jotain tiettyä puoluetta. Lähtökohtaisesti kyseinen tieto ei ole henkilötietoa, sillä tämänkaltaisesta tilastotiedosta ei ole mahdollista selvittää, keitä kyseessä oleva tieto koskee. Tästä huolimatta kyseisen tiedon perusteella on lähtökohtaisesti mahdollista päätellä, että tietyllä alueella asuva äänesti 40 prosentin todennäköisyydellä tiettyä puoluetta. Tällainen tieto voisi periaatteessa olla henkilötietoa, vaikkakin oikeuskirjallisuudessa on katsottu, ettei henkilötiedon käsitettä voitane määritellä näin laajaksi.[209]

Tiettyyn joukkoon tietyllä todennäköisyydellä liittyvän tiedon arvioiminen kuitenkin mahdollisesti muuttuu sitä mukaa, kun tiedon todenperäisyyden todennäköisyys kasvaa. On nimittäin jo vakuuttavammin argumentoitavissa, että tällainen tieto olisi henkilötietoa, kun tieto on totta yli 50 prosentin todennäköisyydellä. Tällöin on mahdollista argumentoida, että tieto on henkilötietoa, sillä tietyllä alueella asuvalla on todennäköisemmin tietty ominaisuus, kuin että hänellä ei kyseistä ominaisuutta olisi. Tilanteen arviointiin vaikuttaa kuitenkin myös arvioitavan henkilöjoukon määrä, sillä on eri asia arvioida viittä henkilöä kymmenen henkilön joukosta kuin 5000 henkilöä 10 000 henkilön joukosta, koska pienemmästä joukosta on mahdollista helpommin tunnistaa yksittäisiä henkilöitä. Siinä tapauksessa, että tieto olisi totta 90 prosentin osalta tietyllä alueella asuvista henkilöistä, on jo periaatteessa mahdollista perustellusti arvioida kyseisen tiedon olevan henkilötietoa, sillä tällöin tietyllä alueella asuvalla on huomattavalla todennäköisyydellä tietty ominaisuus.[210] Tästä huolimatta ei ole mahdollista arvioida varmuudella, ovatko tämänkaltaiset tiedot henkilötietoja missään tilanteissa. Tämä johtuu muun muassa siitä, että tiedon tulkitseminen henkilötiedoksi riippuu aina kontekstisidonnaisista asioista, jonka lisäksi tällainen todennäköisyyksiin perustuva soveltamisalan määrittely ei ole oikeusvarmuuden kannalta ongelmatonta.[211]

2.2.4.2 Tunnistettavuus EUT:n käytännössä: kohtuullisesti toteutettavissa olevien keinojen arviointi

Myös EUT on ottanut kantaa henkilötiedon käsitteen mukaisen tunnistettavuuden tulkintaan vuonna 2016 annetussa ratkaisussaan C-582/14 Breyer, jossa oli kysymys siitä, ovatko dynaamiset IP-osoit-

[209] Korpisaari et al. 2018, s. 55.
[210] Tämän tulkitsemista henkilötiedoksi puoltaisi myös muun muassa se, että jos jokin toimija merkitsisi näiden tietojen perusteella tietoihinsa alueella asuvan henkilön virheellisesti tietyn puolueen äänestäjäksi, kyseinen henkilö ei voisi käyttää TSA 15 ja 16 artiklojen mukaisia oikeuksiaan tarkastaa ja oikaista tietonsa.
[211] Oikeuskirjallisuudessa on myös ylipäänsä kritisoitu tunnistettavuuden konseptia ainoana tietosuojalainsäädännön soveltamisalan määrittäjänä. Ks. esim. Hintze 2018, s. 89.

teet henkilötietodirektiivin määritelmän mukaisia henkilötietoja, kun kaikki tunnistamisen mahdollistavat tiedot eivät olleet yhden tahon hallussa.[212] Tapauksen taustalla oli, että useat Saksan liittovaltion internetsivut tallensivat lokitiedostoihin tietoja internetsivuille kohdistuneesta liikenteestä.[213] Tietojen tallentamisen tavoitteena oli internetsivuille kohdistuneiden hyökkäysten estäminen ja mahdollisten hyökkääjien rikosoikeudelliseen vastuuseen saattaminen, jonka lisäksi internetsivut säilyttivät tietoja lokitiedostoissa senkin jälkeen, kun sivuilla vieraileva poistui sivuilta. Näihin tietoihin sisältyi muun muassa sivuilla käyneen laitteen IP-osoite, vierailun kesto ja vierailun yhteydessä siirretyn tiedon määrä.[214] Tapauksessa useilla Saksan liittovaltion internet-sivuilla vieraillut Breyer oli nostanut paikallisessa saksalaisessa hallintotuomioistuimessa kanteen, jossa hän vaati, että internetsivujen ylläpitäjää kielletään tallentamasta tai antamasta sivullisen tallennettavaksi kyseisillä sivustoilla vierailevan IP-osoitetta internet-sivustoilla käynnin päätyttyä, ellei IP-osoitteen tallentaminen ole tarpeen palvelun käytettävyyden palauttamiseksi häiriötilanteissa.[215] Merkittävässä roolissa Breyerin kanteen taustalla oli se, että hänen internet-yhteyden tarjoajallaan oli hallussaan tiedot, joiden perusteella Breyer oli mahdollista tunnistaa hänen dynaamisen IP-osoitteensa perusteella.

Tapaus päätyi lopulta tarkasteltavaksi Saksan liittovaltion korkeimpaan oikeuteen (Der Bundesgerichtshof, BGH), joka päätyi pohtimaan kysymystä siitä, tuleeko henkilön tunnistettavuutta arvioida ”*objektiivisen*” vai ”*relatiivisen*” kriteerin perusteella.[216] Kysymys tunnistettavuuden määrittelemisestä objektiivisen ja subjektiivisen kriteerin perusteella on oikeuskirjallisuudessa vallinnut debatti tunnistettavuuden arvioinnista.[217] Objektiivisen kriteerin perusteella jo ainoastaan se fakta, että kolmannella osapuolella on hallussaan tiedot, joiden avulla tietty henkilö on tunnistettavissa, tekee rekisterinpitäjän hallussa olevista tiedoista henkilötietoja.[218] Relatiivinen kriteeri tarkoittaa taas sitä, että tiedot ovat henkilötietoja vain sille taholle, joka pystyy tosiasiallisesti tunnistamaan henkilön tiedoista.

[212] C-582/14 Breyer. Vrt. EUT:n ratkaisu C-70/10 Scarlet Extended, jossa oli kyse internetyhteyden tarjoajalle asetettavaksi vaadituista velvollisuuksista internetliikenteen valvomista varten, lausuttiin IP-osoitteiden asemasta henkilötietoina. Ratkaisun mukaan IP-osoitteet ovat henkilötietoja, sillä niiden perusteella on mahdollista tarkasti tunnistaa kyseiset käyttäjät. Ks. myös julkisasiamies Cruz Villalónin ratkaisuehdotus asiassa C-70/10 Scarlet Extended, kohdat 75–77.
[213] C-582/14 Breyer, kohta 14.
[214] Julkisasiamies Sánchez-Bordonan ratkaisuehdotus asiassa C-582/14 Breyer, kohta 23; El Khoury 2017, s. 4. Tällaisia tietoja kutsutaan metatiedoiksi, jotka ovat yksinkertaistettuna ”tietoja tiedoista”. Esimerkiksi kellonaika, jolloin tietystä dynaamisestä IP-osoitteesta vierailtiin internet-sivulla, on verkon metatietoa.
[215] C-582/14 Breyer, kohta 17.
[216] El Khoury 2017, s. 2.
[217] Valinta objektiivisen ja subjektiivisen tulkintakriteerin välillä vaikuttaa muun muassa tietosuojalainsäädännön antaman suojan tosiasialliseen toteutumiseen.
[218] C-582/14 Breyer, kohta 25.

Saksan liittovaltion korkein oikeus totesi *objektiivisen kriteerin* soveltamisen voivan johtaa siihen, että tapauksessa käsiteltävät dynaamiset IP-osoitteet, jotka palveluntarjoaja säilyttää vielä tietyn henkilön poistuttua internet-sivuilta, ovat henkilötietoja. Objektiivisen kriteerin perusteella se on merkityksetöntä, että vain kolmas taho pystyy selvittämään internet-sivuilla vierailleen henkilöllisyyden, ja tiedot ovat tästä huolimatta henkilötietoja myös palveluntarjoajan näkökulmasta.[219] Sen sijaan *relatiivisen kriteerin* soveltamisesta voisi seurata, että dynaaminen IP-osoite olisi henkilötietoa vain Breyerin internet-yhteyden tarjoajan näkökulmasta, sillä vain tämä taho kykenisi tunnistamaan Breyerin kyseisen dynaamisen IP-osoitteen haltijaksi tietyllä hetkellä.[220]

EUT:n ennakkoratkaisun kohteena oli Saksan korkeimman oikeuden kaksi tietosuojalainsäädännön tulkintaa koskevaa kysymystä, joista ensimmäinen on relevantti tutkielman ensimmäisen ja toisen tutkimuskysymyksen näkökulmasta:

> *”Onko [direktiivin 95/46] 2 artiklan a alakohtaa tulkittava siten, että internetprotokollaosoite (IP-osoite), jonka [verkkomediapalvelujen] tarjoaja tallentaa, kun sen internetsivulla käydään, on palveluntarjoajan kannalta henkilötieto jo silloin, kun sivullisella (tässä: yhteyksien tarjoaja) on käytettävissään rekisteröidyn tunnistamiseksi tarvittavat lisätiedot?”*[221]

EUT huomioi perusteluissaan Scarlet Extended ratkaisussa kyseessä olleiden staattisten IP-osoitteiden ja tässä tapauksessa kyseessä olevien dynaamisten IP-osoitteiden välisen eron siten, että näitä kahta tapausta on arvioitava eri tavoin.[222] Staattinen IP-osoite on käytettävään laitteeseen sidottu ja muuttumaton, joten sen perusteella on mahdollista jatkuvasti tunnistaa laite, joka on yhdistetty verkkoon. Uusi dynaaminen IP-osoite taas annetaan laitteelle joka kerta erikseen, kun laite yhdistetään internetiin, eli käytännössä laitteen IP-osoite vaihtuu jokaisen uuden internet-yhteyden myötä. Näin ollen dynaaminen IP-osoite ei ole palveluntarjoajan näkökulmasta *sellaisenaan* yhdistettävissä tiettyyn henkilöön.[223]

---

[219] Zuiderveen Borgesius 2017, s. 131–132.
[220] C-582/14, kohta 25. Ks. myös Zuiderveen Borgesius 2017, s. 131.
[221] C-582/14 Breyer, kohta 30.
[222] C-582/14 Breyer, kohta 36. Vrt. C-70/10 Scarlet Extended, kohdat 50–51. Ks. myös Zuideveen Borgesious 2017, s. 132.
[223] El Khoury 2017, s. 2. Erilaisten IP-osoitteiden välistä eroa kuvaa esimerkki lääkärien takeista. Jos IP-osoitteet olisivat eri värisiä takkeja, joita käytettäisiin sairaalassa lääkärien tunnistamiseksi, staattisen IP-osoitteen käyttäminen tarkoittaisi, että jokainen lääkäri käyttäisi aina samaa takkia. Dynaaminen IP-osoite tarkoittaisi sen sijaan tässä esimerkissä sitä, että aina lääkärin saapuessa sairaalaan, hänelle annettaisiin satunnainen takki kaikkien takkien joukosta. Takki itsessään ei

Tuomioistuin tarkasteli tunnistettavuutta henkilötietodirektiivin 26 johdantokappaleen valossa, joka on myös tietosuojatyöryhmän mukaan keskeinen säännös tunnistettavuuden arvioinnissa.[224] Kyseinen henkilötietodirektiivin 26 johdantokappale vastaa olennaisilta osin TSA:n 26 johdantokappaletta, jonka mukaan tunnistettavuutta arvioitaessa tulee ottaa huomioon ”*kaikki kohtuullisesti toteutettavissa olevat keinot, joita joko rekisterinpitäjä tai joku muu voi kyseisen henkilön tunnistamiseksi käyttää*”.[225] Tietosuojatyöryhmän mukaan kyseistä johdantokappaletta tulisi tulkita siten, että pelkkä teoreettinen mahdollisuus henkilön tunnistamiseen ei riitä täyttämään tunnistettavuuden vaatimusta, vaan tunnistettavuutta tulee arvioida ottamalla huomioon kaikki tunnistamiseen vaikuttavat tekijät.[226] Näin ollen, kun arvioidaan mitä *kohtuullisen todennäköisesti (”reasonably likely”) toteutettavissa olevat keinot* tarkoittavat henkilön tunnistettavuutta arvioitaessa, tulee arviointi suorittaa aina tapauskohtaisesti kunkin tilanteen erityisolosuhteet huomioon ottaen.[227]

Lisäksi tietosuojatyöryhmä korostaa, että rekisterinpitäjän tarkoitus tietojen käsittelylle on merkityksellinen seikka arvioitaessa sitä, onko rekisterinpitäjällä käytettävissään kohtuullisen todennäköisesti toteutettavissa olevat keinot henkilöiden tunnistamiseksi.[228] Lähtökohtana on, että jos tietojen käsittelyn tarkoitus on henkilöiden tunnistaminen, sillä käsittely ei olisi muuten ”perusteltua”, voidaan rekisterinpitäjällä tulkita olevan käytössään tunnistettavuuden mahdollistavat keinot. Tällaista tietojen käsittelyä voisi olla esimerkiksi tietojärjestelmän käyttäjälokien kerääminen. Sen sijaan, jos rekisterinpitäjän tarkoituksena ei ole rekisteröityjen tunnistaminen tiedoista ja rekisterinpitäjä on tietojenkäsittelyn yhteydessä toteuttanut asianmukaiset tekniset ja organisatoriset toimenpiteet henkilöiden tunnistamisen estämiseksi, voi henkilöiden tunnistaminen olla mahdotonta ottaen huomioon kaikki kohtuullisen todennäköisesti käytettävissä olevat keinot.[229]

---

kuitenkaan kertoisi lääkäristä tarpeeksi, jotta hänet voitaisiin tunnistaa, vaan tunnistaminen olisi mahdollista yhdistämällä tieto takin väristä muihin tietoihin.

[224] WP 216, s. 5.

[225] TSA 26 johdantokappale. Niin henkilötietodirektiivin kuin tietosuoja-asetuksenkin 26 johdantokappale käsittelevät henkilön tunnistettavuutta. Käsittelen *tunnistettavuutta* anonyymeihin tietoihin liittyvänä konseptina yksityiskohtaisemmin anonyymejä tietoja ja henkilötietojen anonymisointia käsittelevässä 3 luvussa.

[226] WP 136, s. 15.

[227] Määritelmän ”*kohtuullisen todennäköisesti toteutettavissa olevat*” voidaan katsoa ilmentävän TSA:ssa laajemminkin omaksuttua riskiperusteista suhtautumista henkilötietojen suojaan. Ks. esim. Elliot et al. 2018, s. 206.

[228] WP 136, s. 16.

[229] WP 136, s. 16.

Breyer-ratkaisussa EUT päätyi tulkitsemaan henkilötietodirektiivin 26 johdantokappaletta tietosuojatyöryhmän kantaa vastaavasti, eli siten, että kaiken tunnistettavuuden mahdollistavan tiedon ei tarvitse olla vain yhden tahon saatavilla, jotta henkilö on tunnistettavissa.[230] Tästä johtuen kyseisessä tapauksessa arvioistavaksi tulee, onko palveluntarjoajalla käytettävissään kohtuullisen todennäköisesti sellaiset keinot, joita se voisi käyttää dynaamisen IP-osoitteen yhdistämiseen internet-yhteyden tarjoajan hallussa oleviin tietoihin siten, että kyseisen IP-osoitteen haltija olisi tunnistettavissa.[231]

Tapausta koskevassa julkisasiamiehen ratkaisuehdotuksessa arvioidaan tarkemmin sitä, mitä tarkoitetaan ”*kohtuullisen todennäköisillä keinoilla*”. Julkisasiamies Sánchez-Bordonan mukaan kohtuullisen todennäköisillä keinoilla ei tarkoiteta sellaisia keinoja, jotka olisivat lailla kiellettyjä tai käytännössä mahdottomia.[232] Keinojen katsottaisiin olevan käytännössä mahdottomia, kun henkilön tunnistaminen vaatisi suhteettomia ponnisteluja niin ajallisesti, rahallisesti kuin työvoimallisestikin.[233] Tällaisten keinojen ei tulisi katsoa kuuluvan kohtuullisen todennäköisesti käytössä oleviin keinoihin, sillä henkilön tunnistamisen riski näillä keinoilla olisi merkityksettömän pieni. EUT päätyi ratkaisussaan julkisasiamiehen kanssa samalle kannalle ja lausui, että lailla kiellettyjen keinojen käyttäminen tunnistamiseen ei sisälly kohtuullisen todennäköisten keinojen alaan.[234] Sillä, että tuomioistuin eksplisiittisesti totesi tämän voi olla teoriassa hyvinkin merkittäviä vaikutuksia tietosuojalainsäädännön mukaisen tunnistettavuuden konseptin näkökulmasta.[235] Tästä herää kysymys muun muassa siitä, että jos jäsenvaltion kansallisessa lainsäädännössä säädettäisiin tunnistamisen kieltämisestä tietyissä tapauksissa, voisiko tällaisella säännöksellä rajoittaa kohtuullisten todennäköisten keinojen käytännön soveltumista.[236]

EUT päätyi ratkaisussaan tulkitsemaan direktiivin 95/46/EY 2 artiklan a alakohtaa siten, että dynaaminen IP-osoite on palveluntarjoajaan nähden henkilötieto, jos palveluntarjoajalla on käytettävissään

[230] C-582/14 Breyer, kohta 43; WP 136, s. 13.
[231] C-582/14 Breyer, kohdat 42–45.
[232] Julkisasiamies Sánchez-Bordonan ratkaisuehdotus asiassa C-582/14 Breyer, kohta 68.
[233] C-582/14 Breyer, kohta 46; Julkisasiamies Sánchez-Bordonan ratkaisuehdotus asiassa C-582/14 Breyer, kohta 68.
[234] C-582/14 Breyer, kohta 46. Ks. WP 136, s. 15. Tässä EUT asettui päinvastaiselle kannalle WP 29:n kanssa, sillä tietosuojatyöryhmä toteaa henkilötiedon käsitettä koskevassa lausunnossaan WP 136, että salassapitovelvollisuuden rikkomisesta aiheutuvat organisatoriset häiriöt ovat yksi tekijä, joka tulee ottaa huomioon arvioitaessa kriteeriä ”*kaikki kohtuullisesti toteutettavissa olevat keinot, joita joko rekisterinpitäjä tai joku muu voi kyseisen henkilön tunnistamiseksi käyttää”.*
[235] Toisaalta Zuiderveen Borgesius argumentoi, että EUT:n ratkaisu ei sinänsä ole yllättävä. Ks. Zuiderveen Borgesius 2017, s. 135.
[236] Purtova 2018, s. 65. Tähän liittyen Yhdistyneen Kuningaskunnan uudessa tietosuojalaissa ”*Data Protection Act 2018*” on hyvin mielenkiintoinen luku ”*Offenses relating to personal data*”, jonka 171 kohdan otsikko on ”*Re-identification of de-identified personal data*”. Sen 171(1) kohdan mukaan ”*It is an offence for a person knowingly or recklessly to re-identify information that is de-identified personal data without the consent of the controller responsible for de-identifying the personal data.”.*

lainsäädännössä määrittelyt oikeudelliset keinot, joiden perusteella se voi tunnistaa kyseisen henkilön sellaisten lisätietojen avulla, jotka ovat tämän henkilön internetyhteyden tarjoajan käytettävissä.[237] Tapauksen aikaan voimassa olleen Saksan lainsäädännön mukaan palveluntarjoaja pystyi mahdollisen internet-sivuun kohdistuneen kyberhyökkäyksen yhteydessä ottamaan yhteyttä paikalliseen toimivaltaiseen viranomaiseen, joka kykeni hankkimaan henkilön tunnistamiseen tarvittavat tiedot internetyhteyden tarjoajalta. Tämän seurauksena EUT katsoi, että palveluntarjoajalla oli käytettävissään kohtuullisen todennäköiset keinot henkilön tunnistamiseen, sillä hän pystyi toimivaltaisen viranomaisen ja internetyhteyden tarjoajan avustuksella selvittämään tietyn dynaamisen IP-osoitteen senhetkisen käyttäjän henkilöllisyyden.[238]

EUT:n ratkaisu vastasi keskeiseen kysymykseen dynaamisten IP-osoitteiden luonteesta henkilötietoina, mutta osin ennakkoratkaisua pyytäneen tuomioistuimen tarkkaan rajatuista kysymyksistä johtuen ratkaisu ja sen perustelut kohdistuvat periaatteessa vain hyvin rajattuun yksittäiseen tilanteeseen, eli internetyhteyden tarjoajan hallussa olevien lisätietojen merkitykseen arvioitaessa palveluntarjoajan keräämien dynaamisten IP-osoitteiden luonnetta henkilötietoina.[239] Tästä huolimatta ratkaisusta on mahdollista päätellä useita unionin tuomioistuimen tulkintoja tunnistettavuuden konseptiin liittyen.

Ensinnäkin, kuten julkisasiamies ratkaisuehdotuksessaan toteaa, henkilötietodirektiivin 26 johdantokappaleen mukaisen tunnistettavuuden ”objektiivinen” tai ”absoluuttinen” tulkinta voisi johtaa siihen, että kaikki tieto on henkilötietoa.[240] Objektiivisen kriteerin perusteella tiedosta tulisi henkilötietoa jo ainoastaan sillä perusteella, että jollain kolmannella osapuolella voisi olla hallussaan lisätietoja henkilöstä.[241] Näin tiukka tulkinta ei ole kuitenkaan julkisasiamiehen ratkaisuehdotuksen mukaan perusteltua, sillä ei ole mitään keinoa varmistua sataprosenttisella varmuudella siitä, ettei ole olemassa kolmatta osapuolta, jonka hallussa olisi lisätietoja, joihin tiedot yhdistämällä henkilön voisi tunnistaa.[242] Tästä syystä tunnistettavuuden tulkitseminen objektiivisen kriteerin perusteella ei ole tarkoituksenmukaista, sillä kuten jo tietosuojatyöryhmäkin toteaa, tietosuojasääntöjä ei tulisi soveltaa tilanteisiin, joissa ei ole kyse henkilötietojen käsittelystä tietosuojalainsäädännön merkityksessä.[243]

[237] C-582/14 Breyer, kohta 49.
[238] C-582/14 Breyer, kohta 47.
[239] Julkisasiamies Sánchez-Bordonan ratkaisuehdotus asiassa C-582/14 Breyer, kohdat 61–63.
[240] Julkisasiamies Sánchez-Bordonan ratkaisuehdotus asiassa C-582/14 Breyer, kohta 65.
[241] Toisaalta esim. Zuiderveen Borgesius on argumentoinut, että Breyer-ratkaisun perusteella EUT suosisi tunnistettavuuden arviointia objektiivisen lähestymistavan perusteella, kuten tietosuojatyöryhmäkin. Ks. Zuiderveen Borgesius 2017, s. 135 ja WP 216, s. 9.
[242] Julkisasiamies Sánchez-Bordonan ratkaisuehdotus asiassa C-582/14 Breyer, kohdat 64–65.
[243] WP 136, s. 4.

Toisaalta on kuitenkin ratkaisun perustelujen valossa mahdollista argumentoida, että EUT ei noudattanut tulkinnassaan täysin relatiivista kriteeriä, vaan unionin tuomioistuin otti ratkaisuunsa elementtejä molemmista kriteereistä.[244]

Toiseksi Breyer-tuomion yhteydessä EUT muodosti yhden modernin tietosuojalainsäädännön vaikeimmista konsepteista, kun se *de facto* tunnusti, että on olemassa harmaa alue, jossa tieto voi olla samaan aikaan henkilötietoa sekä anonyymiä tietoa.[245] Tällaisessa tilanteessa kyseessä olevien tietojen asemaa tietosuojalainsäädännön näkökulmasta tulisi Breyer-ratkaisun perusteella tarkastella erikseen objektiivisin kriteerein kunkin toimijan kannalta, jonka hallussa on sellaista tietoa, jota voisi käyttää henkilön tunnistamiseen ottaen huomioon tämän toimijan mahdollisuudet yhdistää tiedot toisiin tietoihin. Toisin sanoen sama tieto voi olla tunnistettavissa olevaa yhdelle rekisterinpitäjälle samaan aikaan, kun se on anonyymiä tietoa toiselle.[246] Lisäksi erityistä huomiota tulisi kiinnittää siihen, mitä keinoja erilaiset toimijat voisivat käyttää luonnollisten henkilöiden tunnistamiseen.

EUT:n ratkaisu osoittaa sen, että henkilötietoja ei ole aina mahdollista tulkita yksiselitteisesti vain joko henkilötiedoiksi tai anonyymeiksi tiedoiksi pelkästään yhden tahon hallussa olevien tietojen perusteella, vaan arvioinnissa tulee ottaa huomioon toimijan tosiasialliset mahdollisuudet henkilön tunnistamiseen yhdistämällä tiedot toisen toimijan hallussa oleviin tietoihin.[247] Tietojen tulkinnassa henkilötiedoiksi, toisin sanoen tunnistettavuuden arvioinnissa, on lopulta kyse siitä, millaisten keinojen katsotaan *kohtuullisen todennäköisesti* mahdollistavan henkilön tunnistamisen. Breyer-ratkaisussa EUT katsoi kohtuullisen todennäköiseksi keinoksi sen, että palveluntarjoajalla oli käytettävissään oikeudelliset keinot, joiden avulla hän pystyi mutkan kautta tunnistamaan Breyerin IP-osoitteen.[248]

Breyer-ratkaisua on mahdollista kritisoida siitä, että kyseisessä ratkaisussa kohtuullisten todennäköisten keinojen alaa voidaan katsoa tulkittavan erittäin laveasti. Henkilötietojen suojaa koskevan lainsäädännön näkökulmasta on perusteltua, että henkilö voi olla tunnistettavissa, vaikka kaikki tunnis-

[244] Ks. El Khoury 2017, s. 5, jonka mukaan henkilötiedon käsitettä tulisi arvioida relatiivisena konseptina. Vrt. kuitenkin Zuiderveen Borgesius 2017, s. 137, jonka mukaan EUT arvioi tunnistettavuutta objektiivisen kriteerin perusteella.
[245] El Khoury 2017, s. 1.
[246] Urgessa 2016, s. 530.
[247] Tietojen tunnistettavuuden arvioimista vaikeuttaa lisäksi se, että nykyaikana tietoja kerätään ja erilaisia tietoja siirrellään ja yhdistellään mitä monimutkaisimmilla tavoilla.
[248] Lisäksi mielenkiintoista on, että yksikään tapausta käsitelleistä tuomioistuimista ei huomioinut sijaintitiedon vaikutusta, sillä useimmiten sekä staattinen että dynaaminen IP-osoite paljastaa käytettävän laitteen sijainnin, ellei käyttäjä käytä erilaisia salauskeinoja, kuten virtuaalista erillisverkkoa (VPN) tai Tor-verkkoa anonymiteettinsä säilyttämiseksi. Toisaalta sijainti IP-osoitteen perusteella on suhteellisen epätarkka.

tamisen mahdollistavat tiedot eivät ole yhden tahon hallussa. Jos kuitenkin kohtuullisen todennäköisesti käytettäviksi keinoksi tulkitaan jo pelkästään poikkeustilanteeseen rajoittuva tiedonsaantimahdollisuus, jonka seurauksena tietoja tulee pitää henkilötietoina, herää kysymys, että tuleeko tunnistettavuutta sittenkin arvioida objektiivisen kriteerin perusteella. Esimerkiksi Breyer-ratkaisussa palveluntarjoajan olisi ollut mahdollista tunnistaa Breyer vain erittäin rajatussa poikkeustilanteessa, jonka toteutumiseen palveluntarjoaja ei olisi voinut juurikaan vaikuttaa. Näin ollen Breyer-ratkaisun perusteella tunnistettavuutta ja kohtuullisen todennäköisesti käytettävissä olevia keinoja arvioitaessa tulee ottaa huomioon myös ainoastaan poikkeustilanteessa käytettävissä olevat keinot, jotka eivät kuitenkaan ole erikseen laissa kiellettyjä.

Breyer-ratkaisun ohella EUT sivusi tunnistettavuutta myös edellä henkilötiedon käsitteen toisen osatekijän yhteydessä käsitellyssä Nowak-ratkaisussa. Unionin tuomioistuin kuitenkin vain totesi ratkaisun yhteydessä sen olevan kiistatonta, että ammatilliseen kokeeseen osallistuja voidaan tunnistaa suoraan tai epäsuorasti koetilanteessa, sillä joko itse koepaperiin tai sen yhteyteen on merkitty hänen nimensä ja henkilönumeronsa.[249] Tässä yhteydessä sillä ei tuomioistuimen mukaan ollut merkitystä, pystyykö koetta tarkastava henkilö tunnistamaan kokeeseen osallistujaa koetta tarkastaessaan vai ei, sillä kaiken tunnistettavuuden mahdollistavan tiedon ei tarvitse olla vain yhden toimijan hallussa.[250]

Tässä kappaleessa yksityiskohtaisesti henkilötiedon käsitteen kannalta käsitelty tunnistettavuuden konsepti liittyy myös olennaisesti anonyymien tietojen ja henkilötietojen anonymisoinnin konseptiin, sillä anonymisoinnilla pyritään poistamaan tunnistettavuus tiedoista siten, että henkilötiedoista tulee anonyymejä tietoja. Näin ollen tarkastelen tunnistettavuutta seikkaperäisesti vielä anonyymien tietojen ja henkilötietojen anonymisoinnin yhteydessä tutkielman 3 luvussa.

2.2.5 Neljäs osatekijä: Luonnollinen henkilö

Neljäs ja samalla viimeinen henkilötiedon käsitteen osatekijä on ”*luonnollinen henkilö”*.[251] Tietosuojalainsäädännön kontekstissa luonnollisen henkilön käsitteellä tarkoitetaan ennen kaikkea sitä, että henkilötietojen suoja koskee ihmisiä heidän kansalaisuudestaan ja asuinpaikastaan riippumatta.[252] Oikeus henkilötietojen suojaan on yleismaallinen ja yksilöiden perusoikeuksia- ja vapauksia tulee

[249] C-434/16 Nowak, kohta 29.
[250] C-434/16 Nowak, kohta 31. Tässä EUT viittasi nimenomaisesti ratkaisun C-582/14 Breyer perustelujen 43 kohtaan.
[251] Myös ilmaisu ”luonnollinen henkilö” on pysynyt TSA:n suomennoksessa samana.
[252] WP 136, s. 21.

kunnioittaa heidän taustastaan riippumatta.[253] Neljännen osatekijän tulkinta aiheuttaa harvoin erimielisyyksiä, vaikka siihenkin liittyy poikkeuksia. Esimerkiksi merkittävä henkilötiedon käsitteeseen liittyvä rajaus on, että luonnollisilla henkilöillä tarkoitetaan vain *eläviä henkilöitä*.[254] Näin ollen eurooppalaisen tietosuojalainsäädännön soveltamisala ulottuu vain elävien henkilöiden henkilötietojen käsittelyyn.

Rekisterinpitäjän tulee tästä TSA:n kuolleiden henkilöiden henkilötietoja koskevasta rajauksesta huolimatta ottaa huomioon se, että kuolleisiin henkilöihin liittyvät tiedot voivat myös liittyä epäsuorasti tunnistettavissa oleviin eläviin henkilöihin ja tällöin tietoja tulee käsitellä tietosuojalainsäädännön vaatimusten mukaisesti. Tällainen tilanne voi olla esimerkiksi kyseessä, jos kuolleella henkilöllä oli perinnöllinen sairaus, joka on automaattisesti myös hänen jälkeläisillään. Näin ollen on rekisterinpitäjän näkökulmasta yksinkertaisesti turvallisempaa käsitellä sekä elävien että kuolleiden henkilötietoja samalla tavalla.[255]

Toisen erikoistilanteen henkilötiedon käsitteen neljännen osatekijän näkökulmasta aiheuttavat syntymättömien lapsien henkilötiedot, joista ei ole sääntelyä EU-oikeudessa. Tästä johtuen on jäsenvaltioiden kansallisesta lainsäädännöstä riippuvaista, että sovelletaanko tietosuojalainsäädäntöä syntymättömiin lapsiin.[256] Käytännössä rekisterinpitäjän voi olla mahdotonta tietää syntymättömistä lapsista, joten tällaiset henkilötiedot tulevat suhteellisen harvoin arvioitavaksi. Kolmas rajaus henkilötiedon käsitteen näkökulmasta aiheutuu oikeushenkilöiden henkilötiedoista. Tietosuojalainsäädäntöä sovelletaan vain luonnollisten henkilöiden henkilötietojen käsittelyyn, joten lähtökohtaisesti henkilötietojen suojaa koskevia säännöksiä ei sovelleta oikeushenkilöihin.[257] Oikeushenkilöihin liittyvät henkilötiedot voivat kuitenkin liittyä epäsuorasti tunnistettavissa oleviin luonnollisiin henkilöihin samalla tavalla kuin kuolleisiin henkilöihin liittyvät tiedot voivat liittyä eläviin henkilöihin. Tiedot voivat esimerkiksi liittyä epäsuorasti yrityksen johdossa oleviin henkilöihin tai yrityksen omistajiin, jolloin tietoja tulisi käsitellä henkilötietoina.[258]

---

[253] WP 136, s. 21. Ks. myös henkilötietodirektiivin 2 johdantokappale.
[254] TSA 27 johdantokappaleen mukaan tietosuoja-asetusta ei sovelleta kuolleisiin henkilöihin. Jäsenvaltioille on kuitenkin jätetty mahdollisuus säätää kansallisissa tietosuojalaeissaan toisin kuolleiden henkilöiden henkilötietojen käsittelystä. Esimerkiksi Suomen tietosuojalakia ei sovelleta kuolleisiin henkilöihin, kun taas Tanskan tietosuojalakia sovelletaan sen 2 §:n 5 kohdan perusteella kuolleisiin henkilöihin, kunnes 10 vuotta on kulunut heidän kuolemastaan.
[255] WP 136, s. 22.
[256] WP 136, s. 23.
[257] Tästä säädetään myös erikseen TSA:ssa, jonka 14 johdantokappaleen mukaan ”*Tämä asetus ei koske oikeushenkilöiden ja erityisesti oikeushenkilön muodossa perustettujen yritysten henkilötietojen käsittelyä, kuten oikeushenkilön nimeä, oikeudellista muotoa ja yhteystietoja”.*
[258] WP 136, s. 23–24.

2.3 Henkilötietojen pseudonymisointi ja pseudonymisoidut tiedot

2.3.1 Henkilötietojen pseudonymisoinnista yleisesti

Henkilötietojen pseudonymisointi[259] ja pseudonymisoidut tiedot ovat oma konseptinsa, joka sijoittuu periaatteessa henkilötietojen ja anonyymien tietojen välimaastossa olevalle harmaalle alueelle.[260] Eurooppalainen tietosuojalainsäädäntö ei eksplisiittisesti tunnustanut pseudonymisointia konseptina ennen tietosuoja-asetusta, sillä henkilötietodirektiivissä ei ollut pseudonymisointia koskevia säännöksiä. Tietosuojatyöryhmä oli kuitenkin arvioinut pseudonymisoituja tietoja henkilötiedon käsitteen näkökulmasta jo vuonna 2007 ja todennut, että henkilöitä on mahdollista tunnistaa pseudonymisoiduista tiedoista epäsuorasti.[261] Näin ollen on tärkeää korostaa, että pseudonymisoidut tiedot ovat tietosuojalainsäädännön systematiikassa edelleen henkilötietoja ja niiden käsittelyyn tulee soveltaa tietosuojasääntöjä.[262]

Pseudonymisoiduilla tiedoilla tarkoitetaan käytännössä sitä, että esimerkiksi nimi ”Mikko Mallikas” ilmaistaan tiedoissa numerosarjalla ”3.593.883”, jolloin henkilöä ei ole mahdollista tunnistaa tiedoista ilman lisätietoja. Siitä huolimatta, että pseudonymisoiduista tiedoista ei ollut aikaisemmin sääntelyä, oli sekä oikeuskirjallisuudessa että käytännön henkilötietojen käsittelyyn liittyvässä toiminnassa henkilötietojen pseudonymisoinnin konsepti ollut esillä jo pitkään ennen tietosuoja-asetusta.[263] TSA:ssa ei määritellä *pseudonymisoituja tietoja*, vaan sen sijaan asetuksen 4(1)(5) artiklassa on *henkilötietojen pseudonymisoinnin* -prosessin määritelmä, jonka mukaan pseudonymisoinnilla tarkoitetaan:

> *”henkilötietojen käsittelemistä siten, että henkilötietoja ei voida enää yhdistää tiettyyn rekisteröityyn käyttämättä lisätietoja, edellyttäen että tällaiset lisätiedot säilytetään*

[259] Pseudonymisointia kutsutaan myös *peitenimillä suojaamiseksi*. Peitenimillä suojaamisesta tietosuojatyöryhmän näkökulmasta ks. WP 216, s. 20–23. Tietosuojatyöryhmän englanninkielisessä lausunnossa pseudonymisointia kutsutaan tosin TSA:n määritelmää vastaavasti termillä ”*pseudonymised data*”.
[260] Tarhonen 2017, s. 10.
[261] WP 136, s. 18–20.
[262] WP 136, s. 18–20; WP 216, s. 20. Tästä huolimatta esimerkiksi TSA:n 28 johdantokappaleen mukaan henkilötietojen pseudonymisointi voi vähentää asianomaisiin rekisteröityihin kohdistuvia riskejä sekä auttaa rekisterinpitäjiä ja henkilötietojen käsittelijöitä noudattamaan tietosuojavelvoitteitaan. Johdantokappaleessa myös erikseen korostetaan, että pseudonymisoinnilla ei ole tarkoitus sulkea pois muita tietosuojatoimenpiteitä.
[263] Oikeuskirjallisuutta henkilötietojen pseudonymisoinnista ennen TSA:ta ks. esim. Zuiderveen Borgesius et al. 2015, s. 2116–2118.

erillään ja niihin sovelletaan teknisiä ja organisatorisia toimenpiteitä, joilla varmistetaan, ettei henkilötietojen yhdistämistä tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön tapahdu.".[264]

Pseudonymisoinnin määritelmän voidaan katsoa muodostuvan kolmesta osasta, jotka kokonaisuutena kuvaavat prosessin, jonka seurauksena henkilötiedoista tulee pseudonymisoituja tietoja.[265] Määritelmästä käy ensinnäkin ilmi, että henkilötietojen pseudonymisointi on TSA 4(1)(2) artiklan mukaista henkilötietojen käsittelyä.[266] Näin ollen pseudonymisointi on sellaista tietojen käsittelyä, joka on TSA:n soveltamisalan piirissä.[267] Toiseksi henkilötietojen pseudonymisoinnin myötä henkilötietoja ei voida enää yhdistää tiettyyn rekisteröityyn käyttämättä lisätietoja. Tämä tarkoittaa käytännössä sitä, että tiedoissa olevat yksilölliset tunnisteet, joiden perusteella henkilöt ovat tunnistettavissa, korvataan sellaisilla tunnisteilla, joiden perusteella henkilöitä ei ole mahdollista tunnistaa.[268] Lisätiedoilla tarkoitetaan tässä yhteydessä useimmiten alkuperäisiä tietoja, joihin pseudonymisoidut tiedot on mahdollista yhdistää tunnistettavuuden mahdollistavalla tavalla. Kolmanneksi henkilötietojen pseudonymisoinnin edellytyksenä on, että lisätiedot tulee säilyttää pseudonymisoiduista tiedoista erillään ja rekisterinpitäjän tulee soveltaa lisätietoihin teknisiä ja organisatorisia toimenpiteitä sen varmistamiseksi, ettei henkilötietoja ole mahdollista yhdistää tunnistettuun tai tunnistettavissa olevaan henkilöön.[269]

Pseudonymisoitujen tietojen asemasta tietosuojalainsäädännön systematiikassa mainitaan vielä erikseen TSA:n 26 johdantokappaleessa, jonka mukaan:

"Pseudonymisoidut henkilötiedot, jotka voitaisiin yhdistää luonnolliseen henkilöön lisätietoja käyttämällä, olisi katsottava tiedoiksi, jotka koskevat tunnistettavissa olevaa henkilöä."[270]

[264] TSA 4(1)(5).
[265] Tarhonen 2017, s. 11.
[266] Henkilötietojen pseudonymisoinnin voidaan katsoa olevan henkilötietojen käsittelyä koskevan TSA:n 4(1)(2) artiklan määritelmän mukaista henkilötietojen "muokkaamista" tai "muuttamista".
[267] Ks. TSA:n 2 artikla asetuksen aineellisesta soveltamisalasta.
[268] WP 136, s. 18.
[269] Teknisistä ja organisatorisista toimenpiteistä säädetään yksityiskohtaisemmin myös muun muassa TSA:n johdantoosan 29, 78, 87 ja 156 johdantokappaleissa.
[270] Tästä käy selkeästi ilmi, että TSA:ssa omaksuttu näkemys pseudonymisoiduista tiedoista vastaa tietosuojatyöryhmän näkemystä.

Tästä syystä pseudonymisoitujen henkilötietojen käsittelyssä tulee edelleen noudattaa TSA:n sääntelyä, kuten tietosuojaperiaatteita ja riskiperusteista lähestymistapaa. Se, että TSA:ssa ei säädellä itse pseudonymisoitujen tietojen asemasta tarkemmin voi periaatteessa hämärtää tällaisten tietojen merkitystä, sillä pseudonymisoitujen tietojen käsittely ei eroa tavanomaisten henkilötietojen käsittelystä. Tästä syystä rekisterinpitäjinä toimivien voi olla vaikea hahmottaa, mitä hyötyä henkilötietojen pseudonymisoinnista itse asiassa on. Toisaalta pseudonymisoitujen tietojen täsmällinen määrittäminen voisi muodostaa TSA:n henkilötietojen ja anonyymien tietojen väliseen kahtiajakoon perustuvaan systematiikkaan sellaisen kolmannen kategorian, joka voisi vaikuttaa keinotekoiselta ja tosiasiallisesti hämärtää tietosuojalainsäädännön soveltamisalaa.[271]

TSA:n mukaisen pseudonymisoinnin tulkinnassa olennaista on sen arvioiminen, milloin ”*henkilötietoja ei voida enää yhdistää tiettyyn rekisteröityyn käyttämättä lisätietoja*”.[272] Rekisteröidyllä tarkoitetaan tässä yhteydessä henkilötiedon käsitteen kolmatta ja neljättä osatekijää: tunnistettu tai tunnistettavissa oleva luonnollinen henkilö. Pseudonymisoinnin arvioinnin kannalta keskeiseksi kysymykseksi näin ollen muodostuu, että milloin henkilö ei ole enää tunnistettavissa tiedoista ilman lisätietoja. Tästä syystä lähtökohtaisesti ainakin suorat tunnisteet, kuten nimet ja henkilötunnisteet tulee korvata pseudonyymeillä, tai poistaa tiedoista. Sen sijaan monimutkaisempaa on arvioida, millaisella tarkkuudella tiedoissa olevat epäsuorat tunnisteet ja niistä muodostettavissa olevat ainutlaatuiset yhdistelmät tulisi henkilötietojen pseudonymisoinnin yhteydessä korvata pseudonyymeillä. Jos tunnistettavuutta arvioitaisiin pseudonymisoinnin yhteydessä samoilla kriteereillä kuin tunnistettavuutta arvioidaan henkilötiedon käsitteen kolmannen osatekijän yhteydessä, olisi henkilötietojen pseudonymisoinnin yhteydessä otettava huomioon muun muassa kaikki kontekstisidonnaiset tekijät, tunnisteiden yhdistelmät ja kohtuullisen todennäköiset keinot.[273]

Oikeuskirjallisuudessa on argumentoitu, että TSA:n vaatimukset täyttävän pseudonymisoinnin kynnystä mahdollisesti alentaisi pseudonymisoinnin määritelmässä käytetty termi ”*tiettyyn rekisteröitytyyn*”.[274] Tämä tarkoittaisi sitä, että henkilötietojen pseudonymisoinnissa tulisi ottaa huomioon lähtökohtaisesti vain sellaiset tiedot, jotka ovat yhdistettävissä vain *yhteen* henkilöön. Näin ollen tiedot,

[271] Ks. esim. ICO 2017a, s. 2, jossa ICO argumentoi, että henkilötiedolle tulisi olla vain yksi määritelmä, ja tästä syystä pseudonymisoituja tietoja ei pitäisi tulkita henkilötiedon käsitteen erillisenä kategoriana.
[272] TSA 4(1)(5). Ks. myös Mourby et al. 2018, s. 223–224.
[273] TSA 4(1)(1); TSA 26 johdantokappale; WP 136; WP 216; C-582/14 Breyer.
[274] Tarhonen 2017, s. 14. Pseudonymisoinnin määritelmä aukeaa paremmin englanniksi, sillä vastaava kohta on TSA:n englanninkielisessä versiossa ”*spesific*”, joka on nyanssisempi käsite kuin ”tietty”.

jotka olisivat yhdistettävissä useampaan henkilöön, eivät olisi merkityksellisiä tietosuojalainsäädännön vaatimukset täyttävän henkilötietojen pseudonymisoinnin arvioinnin näkökulmasta. Tällaisia tietoja olisivat esimerkiksi osoitteet, joissa asuu useampia henkilöitä.[275] Henkilötietojen pseudonymisoinnin määritelmän tulkitseminen tällä tavalla tarkoittaisi käytännössä sitä, että pseudonymisointi täyttäisi TSA:n vaatimukset silloin, kun yksittäisiin henkilöihin liittyvät suorat tunnisteet korvattaisiin pseudonyymeillä.

Pseudonymisointi on ensisijaisesti nähtävissä prosessina, jossa peitetään identiteettejä.[276] Tällaisen prosessin seurauksena rekisteröidystä on mahdollista kerätä lisätietoja ilman, että hänen henkilöllisyyttään selvitetään. Pseudonymisointi on käytännössä mahdollista toteuttaa käyttämällä salaamiseen kaksisuuntaista salausalgoritmia ja säilyttämällä tietojen salausavainta erillään peitenimillä suojatuista tiedoista.[277] Tällaisen prosessin tehokkuus on kuitenkin riippuvainen useasta tekijästä, kuten pseudonymisoitavan aineiston koosta ja sekä käytetystä tekniikasta. Lisäksi oikeaoppisessa pseudonymisoinnissa tulee kiinnittää huomiota siihen, että samaa pseudonyymiä ei käytetä kahden eri henkilön erottamiseen joukosta.[278] Siitä huolimatta, että pseudonymisoidut tiedot liittyvät epäsuorasti tunnistettavissa olevaan luonnolliseen henkilöön, jolloin tietoihin sovelletaan edelleen TSA:n säännöksiä, on henkilötietojen pseudonymisoiminen suositeltavaa.[279] Tällaisten henkilötietojen käsittelyyn liittyy nimittäin tietosuojatyöryhmän mukaan matalampi riski kuin sellaisten tietojen käsittelyyn, jotka ovat suoraan yhdistettävissä tunnistettavissa oleviin henkilöihin. Tämä on myös TSA:ssa omaksuttu sääntelyratkaisu.[280]

Henkilötietojen pseudonymisointi on TSA:ssa tunnustettu prosessi, jonka seurauksena henkilötietojen käsittelyyn liittyvät riskit pienenevät ja kyseisten tietojen turvallisuus paranee[281], sillä esimerkiksi organisaatioon kohdistuneen tietomurron yhteydessä hyökkääjä ei pysty tunnistamaan tiedoissa ole-

[275] Tarhonen 2017, s. 14.
[276] WP 136, s. 18; WP 216, s. 20. Ks. myös ICO 2012, s. 49, jonka mukaan pseudonymisointi on ”The process of distinguishing individuals in a dataset by using a unique identifier which does not reveal their ”real world” identity”.
[277] Kaksisuuntaisella salausalgoritmilla tarkoitetaan salaustekniikkaa, joka mahdollistaa prosessin peruuttamisen tiettyä salausavainta käyttämällä. Ks. luettelo tavanomaisimmista pseudonymisointitekniikoista WP 216, s. 20-21. Tiedot on myös mahdollista salata esimerkiksi käyttämällä yksisuuntaista salausalgoritmia, jolloin tiedoista tulee periaatteessa anonyymejä tietoja.
[278] Saman pseudonyymin käyttäminen kahden eri henkilön erottamiseksi ei ole tarkoituksenmukaista, jos rekisterinpitäjän tarkoituksena on säilyttää tunnistettavuus. Kuitenkin jos tarkoituksena on tehdä yksittäisten henkilöiden tunnistamisesta vaikeampaa, on saman pseudonyymin käyttäminen useampaan henkilöön viittaamiseksi perusteltua.
[279] WP 136, s. 18.
[280] Ks. esimerkiksi TSA 11 ja 12 artiklat.
[281] Tämä vahvistetaan useissa TSA:n johdantokappaleissa, ks. esim. TSA:n johdantokappaleet 75, 78, 85 ja 156.

via henkilöitä ilman toisaalla säilytettyjä lisätietoja. Pseudonymisoinnista seuraavan tietoturvan parantumisen vuoksi TSA:n sääntelyssä rekisterinpitäjiä kannustetaan henkilötietojensa pseudonymisointiin silloin, kun se on tarkoituksenmukaista tietojen käsittelyn näkökulmasta.[282] Henkilötietojen pseudonymisointi mainitaan myös TSA:n 25 artiklassa eksplisiittisesti yhtenä sisäänrakennetun ja oletusarvoisen tietosuojan toteuttamista edistävänä teknisenä ja organisatorisena toimenpiteenä.[283] Lisäksi pseudonymisointi mainitaan asetuksen yleisen edun mukaisia arkistointitarkoituksia taikka tieteellisiä tai historiallisia tutkimustarkoituksia tai tilastollisia tarkoituksia varten tapahtuvaa käsittelyä koskevia suojatoimia ja poikkeuksia koskevassa 89 artiklassa esimerkkinä asianmukaisista suojatoimista, joilla suojellaan rekisteröidyn oikeuksia ja vapauksia.

2.3.2 Henkilötietojen pseudonymisointi ja TSA:n 11 artikla

Henkilötietojen pseudonymisoinnin yhteydessä on syytä huomioida TSA:n 11 artikla, jossa säädetään käsittelystä, joka ei edellytä tunnistamista. Kyseisen artiklan mukaan:

> *”1. Jos tarkoitukset, joihin rekisterinpitäjä käsittelee henkilötietoja, eivät edellytä tai eivät enää edellytä, että rekisterinpitäjä tunnistaa rekisteröidyn, rekisterinpitäjällä ei ole velvollisuutta säilyttää, hankkia tai käsitellä lisätietoja rekisteröidyn tunnistamista varten, jos tämä olisi tarpeen vain tämän asetuksen noudattamiseksi.*
>
> *2. Jos tämän artiklan 1 kohdassa tarkoitetuissa tapauksissa rekisterinpitäjä pystyy osoittamaan, ettei se pysty tunnistamaan rekisteröityä, rekisterinpitäjän on ilmoitettava asiasta rekisteröidylle, jos tämä on mahdollista. Tällaisissa tapauksissa 15–20 artiklaa ei sovelleta, paitsi jos rekisteröity näiden artikloiden mukaisia oikeuksiaan käyttääkseen antaa lisätietoja, joiden avulla hänet voidaan tunnistaa”.*

[282] Ks. TSA 29 johdantokappale, jossa korostetaan pseudonymisointiin tähtäävien kannusteiden luomista rekisterinpitäjille. Johdantokappaleessa lausutaan muun muassa yleisen analyysin mahdollistamisesta, kun henkilötiedot ovat pseudonymisoitu TSA:n sääntelyn näkökulmasta riittävällä tavalla. Pseudonymisointi ei ole kuitenkaan tarkoituksenmukaista esimerkiksi silloin, kun rekisterinpitäjällä on muun lainsäädännön perusteella velvollisuus säilyttää tiedot tunnistettavassa muodossa. Tällainen velvollisuus voi tulla esimerkiksi pankeille rahanpesun estämistä koskevasta lainsäädännöstä.
[283] Henkilötietojen käsittelyn turvallisuuteen liittyen ks. myös TSA 32 artikla, jonka 1a kohdassa mainitaan yhtenä esimerkkinä henkilötietojen pseudonymisointi liittyen henkilötietojen riskiä vastaavan turvallisuustason varmistamiseen.

Lähtökohtaisesti 11 artiklassa säädetään siitä, että jos rekisterinpitäjä ei tunnista rekisteröityä, rekisterinpitäjällä ei ole velvollisuutta ryhtyä toimiin rekisteröidyn tunnistamiseksi, jos käsittely ei edellytä tunnistamista.[284] Tällä on merkittäviä vaikutuksia rekisteröidyn oikeuksien kannalta, sillä tällaisissa tapauksessa TSA:n 15–20 artiklaa[285] ei sovelleta. Kyseisen artiklan käytännön tulkinnan kannalta on keskeistä, että miten *rekisterinpitäjä pystyy osoittamaan, ettei se pysty tunnistamaan rekisteröityä*.[286] Lisäksi jos rekisterinpitäjä siihen pystyy, on sen ilmoitettava tietojen käsittelystä rekisteröidylle. Rekisterinpitäjällä ei kuitenkaan katsottane olevan kovin laajaa velvollisuutta selvittää rekisteröityjen yhteystietoja tällaisissa tilanteissa.[287]

Henkilötietojen pseudonymisoimisen seurauksena rekisterinpitäjä voi päätyä käsittelemään tietoja tavalla, joka ei edellytä tunnistamista. Tällöin rekisterinpitäjän tulisi aluksi pseudonymisoida tiedot TSA 4(1)(5) edellyttämällä tavalla, jonka jälkeen rekisterinpitäjä poistaisi lisätiedot, joiden avulla rekisteröityjä voisi olla vielä mahdollista tunnistaa.[288] Tiedot olisivat kuitenkin edelleen henkilötietoja, sillä yksittäisiä rekisteröityjä voisi olla mahdollista tunnistaa tiedoista esimerkiksi yhdistämällä tällä tavalla pseudonymisoidut tiedot muihin tietoihin.

Periaatteessa TSA:n 11 artikla yhdistettynä henkilötietojen pseudonymisointiin mahdollistaisi sen, että rekisterinpitäjä voisi käsitellä tietoja tunnistamattomassa muodossa ja tällä tavalla suoriutua henkilötietojen käsittelystä kevyemmillä tietosuojavelvoitteilla. Käytännössä merkittävimmäksi kysymykseksi näin ollen muodostuu, että miten rekisterinpitäjä pystyy tehokkaasti osoittamaan, ettei se tunnista rekisteröityä kussakin tilanteessa.[289] Lisäksi jos rekisterinpitäjä joko rekisteröidyn yhteydenoton perusteella, oman toimintansa seurauksena tai sattumalta päätyisi tunnistamaan rekisteröidyn, tulisi tällaiseen henkilötietojen käsittelyyn noudattaa tavanomaisia tietosuojavelvoitteita.[290] Tällöin rekisteröity voisi jälleen käyttää oikeuksiaan täysimääräisesti.

---

[284] Ks. TSA 57 johdantokappale, jonka mukaan rekisterinpitäjää ei saisi velvoittaa hankkimaan lisätietoja rekisteröidyn tunnistamista varten. Tästä huolimatta rekisterinpitäjällä ei saa kieltäytyä vastaanottamasta rekisteröidyn antamia lisätietoja tämän oikeuksien käyttämisen tukemiseksi.
[285] Ks. TSA 15–20 artiklat, joissa säännellään rekisteröidyn oikeudesta saada pääsy tietoihin, oikeudesta tietojen oikaisemiseen, oikeudesta tietojen poistamiseen, oikeudesta käsittelyn rajoittamiseen, henkilötietojen oikaisua tai poistoa taikka käsittelyn rajoitusta koskevasta ilmoitusvelvollisuudesta sekä oikeudesta siirtää tiedot järjestelmästä toiseen.
[286] Esimerkiksi tutkielman johdannossa tarkastellussa esimerkissä Lontoon metron käyttäjistä kerätyistä tiedoista TfL:n projektista vastaava henkilö oli todennut, ettei TfL pysty anonymisoinnin vuoksi vastaamaan rekisteröityjen tietopyyntöihin. Tästä näkökulmasta on mahdollista pohtia, että oliko Lontoon metron anonymisoinnissa kyse pikemminkin TSA:n 11 artiklan mukaisesta menettelystä.
[287] Korpisaari et al. 2018, s. 169.
[288] Hintze 2018, s. 92.
[289] Muun muassa Purtova 2018, s. 79 on todennut, että 11 artiklan hyödyntämisen ongelmaksi muodostuu, miten sen tulkinta rajataan koskemaan vain tilanteita, joissa sen käyttäminen on perusteltua.
[290] Data Protection Handbook 2018, s. 94.

2.4 Päätelmiä henkilötiedon käsitteestä

Kuten edeltä käy ilmi, muodostuu henkilötiedon käsitteen määritelmä neljästä toisiinsa vahvasti linkittyneestä osatekijästä, jotka ovat ”*kaikenlaiset tiedot*”, ”*liittyvä*”, ”*tunnistettu tai tunnistettavissa oleva*” ja ”*luonnollinen henkilö*”. Näin ollen henkilötiedon määritelmä on erittäin laaja, muttei kuitenkaan rajaton, sillä esimerkiksi EUT:n Breyer-ratkaisussa dynaamista osoitetta sellaisenaan ei pidetty henkilötietona.[291] Lisäksi tietosuojatyöryhmä korostaa henkilötiedon käsitettä koskevassa lausunnossaan, ettei ole toivottavaa soveltaa tietosuojasääntöjä tilanteisiin, joihin niitä ei ole tarkoitettu. Toisaalta lähes samaan hengenvetoon tietosuojatyöryhmä lausuu, ettei henkilötiedon käsitteen määritelmän tulkintaa tule rajoittaa aiheettomasti.[292] Tästä on pääteltävissä, että tietojen tulkitseminen henkilötiedoiksi on vahvasti kontekstisidonnaista ja tietojen käsittelyn yhteydessä on perusteltua aina huomioida käsittelyn vaikutukset luonnollisten henkilöiden oikeuksien kannalta.

Henkilötiedon käsitettä on tarkoituksenmukaista arvioida neljän tietosuojatyöryhmän henkilötiedon käsitteestä muodostaman osatekijän näkökulmasta. Näiden osatekijöiden valossa sen arvioiminen, onko kussakin tapauksessa todella kyse henkilötiedoista, muodostuu seuraavista kysymyksistä: 1) Ovatko käsiteltävät tiedot luonteeltaan sellaisia tietoja, jotka voivat olla henkilötietoja? 2) Liittyvätkö tiedot tiettyyn henkilöön joko sisältönsä, tarkoituksensa tai tuloksensa (EUT:n Nowak-ratkaisun valossa vaikutuksensa) vuoksi? 3) Onko yksittäinen henkilö tunnistettavissa tiedoista? 4) Koskevatko tiedot eläviä luonnollisia henkilöitä?

Vastaus ensimmäiseen kysymykseen on lähes poikkeuksetta kyllä, sillä sekä tietosuojatyöryhmä että EUT ovat arvioineet henkilötiedon käsitteen ensimmäistä osatekijää erittäin laajasti, ja kaikki objektiiviset ja subjektiiviset tiedot voivat olla henkilötietoja.[293] Eräs harvoista poikkeuksista, jolloin tieto ei ole ensimmäisen osatekijän valossa henkilötietoa on, että tieto on ainoastaan tietyn henkilön muistissa. Toiseen kysymykseen vastaaminen on kuitenkin jo monimutkaisempaa, sillä muun muassa EUT ei ole ollut toisen osatekijän tulkinnassa yhdenmukainen.[294] Tietosuojatyöryhmän mukaan tiedot lähtökohtaisesti liittyvät tiettyyn henkilöön, kun ne kertovat hänestä.

[291] C-582/14 Breyer, kohta 38.
[292] WP 136, s. 5.
[293] WP 136, s. 6; C-434/16 Nowak, kohta 34.
[294] Vrt. keskenään ratkaisuja C-141/12 YS ym. ja C-434/16 Nowak.

Tietojen sisältö on käytännössä yksinkertaisin edellytys sille, että tiedot liittyvät tiettyyn henkilöön: tämä vastaa yleiskielen määritelmää yksilöön liittyvän tiedon sisällöstä. Tiedot sen sijaan liittyvät yksilöön niiden tarkoituksensa vuoksi silloin, kun tietojen tarkoituksena on yksilön arvioiminen taikka tietyllä tavalla kohtelu tai häneen jollain tavalla vaikuttamisensa. EUT:n YS ym.-ratkaisussa oleskeluvaluvan hakijan oleskelulupapäätöstä koskevan oikeudellisen arvioinnin ei arvioitu liittyvän henkilöön riittävällä tavalla, sillä tuomioistuimen perustelujen mukaan kyse oli ainoastaan oikeussääntöjen abstraktista tulkinnasta.[295] Tietosuojatyöryhmän mukaan kyseessä oleva oikeudellinen arviointi olisi kuitenkin mitä todennäköisemmin liittynyt oleskeluluvan hakijaan henkilötiedon käsitteen toisen osatekijän mukaisesti, sillä tämänkaltaisella arviolla tosiasiallisesti arvioidaan henkilön oikeutta oleskelulupaan.

Tietojen tulos tai vaikutus on taas käsillä, jos tietojen käsittelyn tuloksena henkilöä tullaan todennäköisesti kohtelemaan eri tavalla kuin jos tietoja ei olisi käsitelty. Näin ollen tietosuojatyöryhmän tulkinnan perusteella tiedot voivat liittyä henkilöön henkilötiedon käsitteen toisen osatekijän näkökulmasta jo ainoastaan siitä syystä, että niiden käsittelyn seurauksena henkilöä mahdollisesti kohdellaan toisenlaisessa tavalla, kuin kyseisiä tietoja ei olisi käsitelty. Huomionarvoista on, että tietosuojatyöryhmän mukaan tulostekijän arvioinnissa tulee ottaa huomioon se, jos henkilöä kohdellaan ”todennäköisesti” eri tavalla, erotuksena tunnistettavuuden arvioinnissa vaikuttavaan ”kohtuullisen todennäköisten keinojen” arviointiin.[296] Tästä näkökulmasta vaikuttaa siltä, että tietosuojatyöryhmän mukaan tiedot liittyvät tiettyyn henkilöön alhaisemmalla kynnyksellä kuin henkilö olisi tiedoista tunnistettavissa. EUT ei huomioinut tarkoitustekijän ohella tulostekijää YS ym.-ratkaisussaan, sillä edellä mainitun oikeudellisen arvioinnin tuloksena mitä ilmeisimmin on, että oleskeluluvan hakijaa kohdellaan eri tavalla, sillä tietojen käsittelyn tulos määrittää sen, onko henkilö oikeutettu oleskelulupaan. EUT kuitenkin päätyi tarkoitus- ja tulostekijän osalta tietosuojatyöryhmän kanssa vastaavaan tulkintaan tuoreimmassa henkilötiedon käsitettä koskevassa Nowak-ratkaisussaan, jossa EUT nimenomaisesti totesi ratkaisun poikkeavan YS ym.-ratkaisusta.[297] Näin ollen tietojen katsotaan henkilötiedon käsitteen toisen osatekijän mukaan liittyvän tiettyyn henkilöön, kun tietojen käsittelyn seurauksena henkilöä kohdellaan todennäköisesti eri tavalla.

[295] C-141/12 YS ym., kohta 40.
[296] WP 136, s. 10; Purtova 2018, s. 54–55.
[297] C-434/16 Nowak, kohta 56; C-141/12 YS. ym., kohdat 45–46.

Henkilötiedon kolmanteen osatekijään liittyvä kolmas kysymys on samaan aikaan sekä tutkielman tutkimuskysymysten näkökulmasta tärkein että henkilötiedon käsitteen määritelmän kannalta merkittävin, sillä tunnistettavuus määrittää useimmiten käytännössä sen, ovatko kyseessä olevat tiedot henkilötietoja.[298] Tunnistettavuuden täsmällinen määrittäminen on kuitenkin TSA:n 26 johdantokappaleen, tietosuojatyöryhmän tulkintakannanottojen ja EUT:n soveltuvan ratkaisukäytännön valossa haasteellista. Tämä johtuu toisaalta siitä, että kohtuullisen todennäköisesti käytettävissä olevien keinojen arviointi on hyvin kontekstisidonnaista, toisaalta siitä, että kyseisiä keinoja on tulkittu erittäin laajasti. Muun muassa EUT:n Breyer-ratkaisussa omaksuman tulkinnan mukaan kohtuullisen todennäköisesti käytettävissä oleviksi keinoiksi katsottiin jo pelkkä poikkeukselliseen erityistilanteeseen rajoittuva tiedonsaantioikeus, jonka seurauksena tiedoista tuli henkilötietoja. Samalla EUT kuitenkin lausui yhden kyseisten keinojen tulkintaa rajoittavan seikan: tällaisiksi keinoiksi ei tulkita erikseen laissa kiellettyjä keinoja.[299]

Lisäksi merkittävästi tunnistettavuuden konseptia laajentava ja sen arvioinnissa huomioon otettava seikka on, ettei tunnistettavuuden mahdollistavien muiden tietojen tarvitse olla rekisterinpitäjän tai ylipäänsä vain yhden tahon hallussa.[300] Tämä on välttämätöntä henkilötietojen suojan näkökulmasta, sillä muuten organisaatiot voisivat kiertää tietosuojalainsäädännön vaatimukset vetoamalla siihen, että heidän hallussaan ei ole tunnisteellista tietoa, vaikka tällaisen tiedon hankkiminen olisi organisaatioille yksinkertaista. Näin ollen tunnistettavuuden arvioinnissa tulee ottaa huomioon kaikki lisätiedot, jotka voivat mahdollistaa henkilön tunnistamisen tietoja yhdistelemällä. Kuitenkaan pelkkä teoreettinen mahdollisuus, että kolmannen osapuolen hallussa voisi olla tunnistamisen mahdollistavia tietoja ei vielä riitä täyttämään tunnistettavuuden edellytyksiä, joten tunnistettavuutta ei tule tulkita ainoastaan absoluuttisen kriteerin perusteella.[301]

Vastaus neljänteen kysymykseen on periaatteessa hyvin yksinkertainen: tietosuojalainsäädäntöä sovelletaan luonnollisiin henkilöihin heidän taustastaan, kansalaisuudestaan ja asuinpaikastaan riippumatta.[302] Tämän osatekijän soveltumista rajoittaa kuitenkin se, että TSA:ta sovelletaan vain eläviin

[298] WP 136, s. 4.
[299] C-582/14 Breyer, kohta 46.
[300] C-582/14 Breyer, kohta 43. Ks. tähän liittyen Quinn – Quinn 2018, s. 1002, jossa kirjoittajat toteavat, että big datan käytön hyödyntämisen myötä etenkin geenitiedoista on nykyisin mahdollista tunnistaa yksittäisiä henkilöitä ennennäkemättömän tehokkaasti yhdistämällä tiedot aikaisemmin julkaistuissa tutkimuksissa oleviin tietoihin.
[301] C-582/14, kohta 46; Julkisasiamies Sánchez-Bordonan ratkaisuehdotus asiassa C-582/14 Breyer, kohta 68.
[302] Huomioiden kuitenkin TSA:n muut soveltamisalan rajaukset, kuten kotitalouskäytön ja maantieteellisen soveltamisalan.

luonnollisiin henkilöihin.[303] Näin ollen kuolleiden henkilöiden henkilötietojen käsittelyyn ei sovelleta tietosuojalainsäädäntöä.[304]

Yhteenvetona henkilötiedon käsitteen voidaan todeta olevan erittäin laaja, joustava ja teknologiseen kontekstiin mukautuva.[305] Käsite on laaja, jotta luonnollisten henkilöiden perusoikeudet, etenkin oikeus henkilötietojen suojaan henkilötietojen käsittelyssä voidaan turvata. Tämän lisäksi käsite on joustava, ettei tietosuojalainsäädännön kiertäminen olisi mahdollista muun muassa vetoamalla siihen, että tiedot eivät liity henkilöihin, koska ne eivät vaikuta heidän asemaansa.[306] Lopuksi käsite on teknologiseen kontekstiin mukautuva, eli teknologianeutraali, jotta esimerkiksi tunnistettavuuden tulkinnassa voidaan ottaa huomioon myös uudet teknologian mahdollistamat tunnistuskeinot, joista emme välttämättä vielä ole edes tietoisia.[307]

Kuten edellä kappaleessa 2.1.2 mainittiin, TSA:lla on sen 1 artiklan perusteella kaksi tavoitetta: henkilötietojen suojeleminen ja henkilötietojen vapaa liikkuvuus. Yksi asetuksen keskeisimmistä ongelmista on, että nämä asetuksen tavoitteet ovat ristiriidassa keskenään, ja tulkinnassa tulisi tasapainotella perusoikeuksien suojan ja taloudellisten intressien huomioisen välillä. Tästä tasapainottelusta seurannut EUT:n ratkaisukäytäntö on kuitenkin 2010-luvulla korostanut henkilötietojen suojan merkitystä sillä seurauksella, että henkilötietojen vapaa liikkuvuus on vaikuttanut aika ajoin lähes unohdetulta tavoitteelta. Muun muassa julkisasiamies Bobek on tähän liittyen todennut, että EUT:n äskettäistä oikeuskäytäntöä tietosuojan alalla on leimannut toive varmistaa henkilötietojen tehokas suoja.[308] Henkilötietojen suojan korostumisen myötä EUT on päätynyt tietosuojalainsäädännön tulkinnallaan laajentamaan henkilötiedon käsitteen tulkintaa, kun yhä useammin yksittäinen tiedon on katsottu liittyvän tunnistettavissa olevaan luonnolliseen henkilöön.

Arvioitaessa EUT:n henkilötiedon käsitteen tulkintaa koskevia ratkaisuja, on tärkeää pohtia millaisia seurauksia päinvastaisilla ratkaisuilla olisi ollut. Tämä on samalla laajempi kysymys EUT:n henkilötiedon käsitettä koskeviin ratkaisuihin liittyen, sillä nämä ratkaisut käytännössä joko rajaavat tai laajentavat koko eurooppalaisen tietosuojalainsäädännön soveltamisalaa. Jos yksittäisessä tapauksessa tietojen ei tulkita olevan henkilötietoja joko tunnistettavuuden puuttumisen seurauksena, tai koska ne

[303] Tästä on kuitenkin mahdollista säätää kansallisesti toisin, ks. TSA 27 johdantokappale.
[304] WP 136, s. 25; TSA 27 johdantokappale.
[305] Purtova 2018, s. 44.
[306] WP 136, s. 11.
[307] TSA 26 johdantokappale; Ks. myös Purtova 2018, s. 44.
[308] Julkisasiamies Bobekin ratkaisuehdotus asiassa C-40/17 FashionID, kohta 72.

eivät liity tunnistettuun tai tunnistettavissa olevaan henkilöön, voi tällä olla merkittäviä vaikutuksia henkilötietojen suojan kannalta. Tämä voi myös vaikuttaa osaltaan siihen, että EUT pyrkii lähtökohtaisesti mieluummin tulkitsemaan yksittäiset tiedot henkilötiedoiksi, kuin anonyymeiksi tiedoiksi.

3 Anonyymit tiedot ja henkilötietojen anonymisointi

3.1 Anonyymit tiedot eurooppalaisessa tietosuojalainsäädännössä

3.1.1 Henkilötietojen ja anonyymien tietojen välinen suhde

Tietosuojalainsäädäntö on järjestelmä, jonka soveltuminen edellyttää, että käsiteltävät tiedot tulkitaan henkilötiedoiksi. Tämän vuoksi tietosuojalainsäädännön näkökulmasta kaikki järjestelmän, toisin sanoen soveltamisalan, ulkopuolelle jäävät tiedot – tiedot, jotka eivät ole henkilötietoja – on katsottava anonyymeiksi tiedoiksi. Kun eurooppalaisessa tietosuojalainsäädännössä kehitettiin ensimmäistä kertaa henkilötiedon käsite, muodostuikin samalla implisiittisesti myös anonyymien tiedon käsite. Näin ollen eurooppalaisen henkilötietojen suojaa koskevan lainsäädännön systematiikan voidaan katsoa perustuvan henkilötietojen ja anonyymien tietojen väliselle jaottelulle. Tulkinta tietojen henkilötietoudesta ratkaisee sen, soveltuvatko henkilötietojen suojaa koskevat säännökset lainkaan. Tietosuojalainsäädäntö on nähtävissä suljettuna järjestelmänä, jonka porttina henkilötiedon käsite toimii.

Tällaisessa järjestelmässä olisi tarkoituksenmukaista, että jaottelu henkilötietojen ja anonyymien tietojen välillä olisi mahdollisimman selkeä ja helposti ymmärrettävissä, kun henkilötiedon käsitteen tulkinta määrittää koko tietosuojalainsäädännön soveltumisen.[309] Vasta sitten, kun tiedot katsotaan henkilötiedoiksi, tulee pohdittavaksi muun muassa tietosuojaperiaatteiden toteutuminen henkilötietojen käsittelyssä.[310] Käytännössä erottelu henkilötietojen ja anonyymien tietojen välillä voi olla kuitenkin erittäin haastavaa. Tämä johtuu muun muassa henkilötiedon käsitteen avoimen laajaksi kirjoitetusta määritelmästä[311], tietosuojatyöryhmän ja EUT:n henkilötiedon käsitteen todella laajasta tulkinnasta[312] sekä tunnistamisen mahdollistavien teknologioiden nopeasta kehityksestä.[313]

Henkilötietojen ja anonyymien tietojen välien rajanveto on modernissa tiedon hyödyntämiseen perustuvassa maailmassa erittäin tärkeää, sillä esimerkiksi useat eurooppalaiset julkisen sektorin toimijat pyrkivät muodostamaan hallussaan olevista henkilötiedoista anonyymejä tietoja anonymisoinnin

[309] Stalla-Bourdillon – Knight 2016, s. 320–321. Toisaalta oikeuskirjallisuudessa on myös argumentoitu, että henkilötietojen ja anonyymien tietojen välinen rajanveto on odotettua ongelmattomampaa.
[310] Tietosuoja-asetus 5 artikla. Tietosuojaperiaatteista tarkemmin, ks. esim. de Hert et al. 2013.
[311] Tietosuoja-asetus 4(1)(1) artikla.
[312] Ks. esim. Edellä mainitut tietosuojatyöryhmän lausunto WP 136 ja EUT:n ratkaisut C-582/14 Breyer ja C-343/16 Nowak.
[313] Purtovan 2018, s. 41 mukaan informaatioteknologian kehityksen seurauksena kaikesta ympärillämme tapahtuvasta kerätään pian tietoa tavalla, josta yksilöt on mahdollista tunnistaa.

keinoin, jotta ne voisivat muun muassa julkaista tällaista tietoa avoimena datana.[314] Samalla erilaisten yksityisen sektorin toimijoiden intressissä on muodostaa henkilötietovarannoistaan anonymisoinnilla sellaisia anonyymejä tietoja, joita ne voisivat analysoida data-analytiikan menetelmillä vapaasti ilman tietosuojalainsäädännön rajoituksia.[315] Tosiasiallisesti kuitenkin tietosuojalainsäädännön vaatimukset täyttävän anonymisoinnin toteuttaminen on erittäin vaikeaa, jos pyrkimyksenä on säilyttää tietojen hyödynnettävyys myöhempiin käyttötarkoituksiin.[316]

3.1.2 Anonyymien tietojen ja anonymisoinnin oikeudellinen määritelmä

Anonyymien tietojen määritelmä tietosuoja-asetuksessa on negatiivinen määritelmä.[317] Tämä tarkoittaa sitä, että tiedot, jotka eivät ole henkilötietoja, ovat anonyymejä tietoja. TSA:ssa ei ole anonyymien tietojen määritelmää artiklatasolla, vaan sen sijaan anonyymeista tiedoista säädetään ainoastaan asetuksen johdanto-osan 26 johdantokappaleessa, jonka mukaan:

> *”Tietosuojaperiaatteita ei […] pitäisi soveltaa anonyymeihin tietoihin eli tietoihin, jotka eivät liity tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön, tai henkilötietoihin, joiden tunnistettavuus on poistettu siten, ettei rekisteröidyn tunnistaminen ole tai ei ole enää mahdollista. Tämä asetus ei tämän vuoksi koske tällaisten anonyymien, muun muassa tilasto- tai tutkimustarkoituksen varten käytettävien tietojen käsittelyä”.*[318]

Tässä johdantokappaleessa eksplisiittisesti ilmaistaan se, että TSA:ta ei sovelleta anonyymeihin tietoihin.[319] Eurooppalainen lainsäätäjä on selkeästi tunnistanut anonyymien tietojen konseptin, mutta tietoisesti jättänyt sen syvällisemmän määrittelemisen tietosuojalainsäädännön ulkopuolelle.[320] Joh-

[314] Zuiderveen Borgesius et al. 2015, s. 2078–2079; Esayas 2015, s. 3. Ks. myös UKAN 2016, s. 7.
[315] Leonard 2014, s. 60. Anonymisoidun tiedon hyödyntäminen on tarkoituksenmukaista etenkin big data –analytiikassa.
[316] ENISA 2015, s. 27.
[317] Tamò Larrieux 2018, s. 92.
[318] TSA 26 johdantokappale.
[319] EU-oikeudessa on kuitenkin sääntelyä, jota sovelletaan kaikenlaisen tiedon käsittelyyn, oli se henkilötietoa tai ei. Ks. esim. sähköisen viestinnän tietosuojadirektiivin mukaisia tietoja koskevat säännökset.
[320] On kiinnostavaa, että eurooppalainen lainsäätäjä oli sisällyttänyt anonyymin tiedon ja anonymisoinnin määritelmän ensimmäiseen ehdotukseen henkilötietodirektiiviksi, mutta päätynyt poistamaan sen direktiivin valmistelun myöhemmässä vaiheessa. Ks. Euroopan komissio 1990, s. 19, jossa komissio on määritellyt käsitteen ”*depersonalize*”. Komission ehdotuksen mukaan ”*’Depersonalize’. This concept is designed to permit the exclusion from the scope of certain provisions of the directive of data which are no longer identifiable. An item of data can be regarded as depersonalized even if it could theoretically be repersonalized with the help of disproportionate technical and financial resources*”. Kaksi vuotta myöhemmin komissio kommentoi täydennetyssä ehdotuksessa henkilötietodirektiiviksi Euroopan komissio 1992, s. 10

dantokappaleesta on kuitenkin pääteltävissä sekä anonyymien tietojen että henkilötietojen anonymisoinnin oikeudelliset määritelmät. Sen perusteella *anonyymeillä tiedoilla* tarkoitetaan sellaisia tietoja, jotka eivät liity tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön. Johdantokappaleesta johdettavissa oleva henkilötietojen anonymisoinnin määritelmä liittyy taas erottamattomasti anonyymien tietojen määritelmään. *Henkilötietojen anonymisointi* tarkoittaa prosessia, jonka seurauksena henkilötiedoista poistetaan tunnistettavuus siten, ettei rekisteröidyn tunnistaminen ole enää mahdollista.[321] Näin ollen johdantokappaleen mukaan edellä 2.2.3 kappaleessa yksityiskohtaisesti käsitelty *tunnistettavuus* määrittää sen, onko jokin tieto anonyymiä tietoa vai henkilötietoa.[322]

TSA:n 26 johdantokappaleessa edelleen tarkennetaan tunnistettavuuden määrittämistä ja täsmennetään sitä, mitä on otettava huomioon arvioitaessa luonnollisen henkilön tunnistettavuutta. Johdantokappaleen mukaan:

> *”Tietosuojaperiaatteita olisi sovellettava kaikkiin tietoihin, jotka koskevat tunnistettua tai tunnistettavissa olevaa luonnollista henkilöä. […] Jotta voidaan määrittää, onko luonnollinen henkilö tunnistettavissa, olisi otettava huomioon kaikki keinot, joita joko rekisterinpitäjä tai muu henkilö voi kohtuullisen todennäköisesti käyttää mainitun luonnollisen henkilön tunnistamiseen suoraan tai välillisesti, kuten kyseisen henkilön erottaminen muista. Jotta voidaan varmistaa, voidaanko keinoja kohtuullisen todennäköisesti käyttää luonnollisen henkilön tunnistamiseen, olisi otettava huomioon kaikki objektiiviset tekijät, kuten tunnistamisesta aiheutuvat kulut ja tunnistamiseen tarvittava aika sekä käsittelyajankohtana käytettävissä oleva teknologia ja tekninen kehitys”.*[323]

---

kyseistä määritelmää uudelleen. Komission mukaan: ”*Depersonalized" data are not defined: the term is not used in the Directive. This means that whether or not data are depersonalized no longer depends on the cost of determining the data subject's identity (amendment No 13). However, in the specific case where data are compiled in the form of statistics, it has been considered appropriate to state that they cannot be considered to be personal data where the data subjects can no longer reasonably be identified*”.

[321] Sähköisen viestinnän tietosuojadirektiivissä säädetään myös anonyymeistä tiedoista ja anonymisoinnista TSA:n sääntelyä vastaavalla tavalla. Sen 26 johdantokappaleen mukaan “*Traffic data used for marketing communications services or for the provision of value added services should also be erased or made anonymous after the provision of the service*”. Kyseisen direktiivin suomenkielisessä versiossa 26 kohdassa säädetään “nimettömäksi tekemisestä” anonymisoinnin sijaan, joten englanninkielinen versio on anonymisoinnin tulkinnan kannalta selkeämpi. Ks. myös saman direktiivin 6(1) ja 9(1) artiklat, joissa säädetään vaatimuksista anonymisoida tilaajia ja käyttäjiä koskevat liikenne- ja paikkatiedot, kun tiedot ovat tarpeellisia tai käyttäjät ovat antaneet käsittelyyn suostumuksensa.

[322] Tunnistettavuuden konseptin arvioinnista muun muassa tietosuojatyöryhmän näkökulmasta, ks. tutkielman jakso 2.2.4.

[323] Vertailuksi aikaisemman henkilötietodirektiivin 26 johdantokappaleen mukaan ”*Tietosuojaa koskevia periaatteita on sovellettava kaikkiin tunnistettua tai tunnistettavissa olevaa henkilöä koskeviin tietoihin; sen määrittämiseksi, onko henkilö tunnistettavissa, olisi otettava huomioon kaikki kohtuullisesti toteutettavissa olevat keinot, joita joko rekisterinpitäjä tai joku muu voi kyseisen henkilön tunnistamiseksi käyttää; tietosuojaa koskevia periaatteita ei sovelleta tietoihin, jotka on tehty anonyymeiksi siten, ettei rekisteröity enää ole tunnistettavissa*;”. Vertailtaessa henkilötietodirektiivin ja tietosuoja-asetuksen 26 johdantokappaleita on huomattavissa, että anonyymin tiedon määritelmä on pysynyt jokseenkin sa-

Näin ollen TSA:n 26 johdantokappaleen valossa tunnistamista tulee arvioida objektiivisin kriteerein, joita ovat kohdassa *esimerkkeinä* mainitut 1*) tunnistamisesta aiheutuvat kulut*, 2) *tunnistamiseen tarvittava aika*, 3) *käsittelyajankohtana käytettävissä oleva teknologia* ja 4) *tekninen kehitys*. Johdantokappaleessa oleva lista erilaisista tunnistettavuuden arvioinnin kriteereistä ei ole tyhjentävä, sillä kyseisen kohdan mukaan arvioinnissa on otettava huomioon kaikki objektiiviset tekijät, *kuten* kohdassa lausutut neljä kriteeriä. Näin ollen johdantokappale jättää avoimeksi sen, millä muilla kriteereillä tunnistettavuutta tulisi arvioida. Käytännössä lienee kuitenkin perusteltua arvioida 26 johdantokappaletta ottaen huomioon kaikki muut, paitsi laittomat keinot, joita rekisterinpitäjä tai kolmas voi käyttää henkilöiden tunnistamiseksi.[324] Joka tapauksessa, jos tunnistaminen ei objektiivisten kriteerien valossa ole *kohtuullisen todennäköisesti* mahdollista, tulisi tiedot tulkita anonyymeiksi tiedoiksi, jotka eivät liity tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön.[325]

Käytännössä TSA:n 26 johdantokappale edellyttää sitä, että henkilötietojen anonymisoinnin toteuttamiseksi tiedoista on poistettava riittävästi elementtejä siten, ettei yksikään taho enää kykene tunnistamaan luonnollisia henkilöitä kyseisistä tiedoista.[326] Erityisen tärkeä kriteeri anonymisointiin liittyen on käsittelyn peruuttamattomuus. Rekisterinpitäjä tai kolmas ei saa pystyä peruuttamaan anonymisointia, sillä jos tämä on mahdollista, tietoja on käsiteltävä henkilötietoina.[327] Tietosuojalainsäädännön vaatimukset täyttävän henkilötietojen anonymisoinnin arvioinnin kannalta keskeinen sananpari on sama kuin tunnistettavuuden arvioinnissa: ”*kohtuullisen todennäköisesti”*. Tästä johtuen henkilöiden tunnistamisen ei täydy olla objektiivisesti arvioiden mahdotonta, jotta tiedoista tulee anonymisoinnin myötä anonyymejä tietoja.[328] Lisäksi on huomioitava, että henkilötietojen anonymisoinnin yhteydessä rekisterinpitäjän, kolmannen tai tietojen vastaanottajan aikomuksilla ei ole merkitystä tietosuojalainsäädännön soveltumisen arvioinnissa ja tietosuojaa koskevaa lainsäädäntöä tulee soveltaa lähtökohtaisesti aina, kun tiedoista voidaan tunnistaa yksittäisiä henkilöitä.[329]

---

mana ja TSA:n 26 johdantokappale vain tarkensi tunnistettavuuden arvioinnin määritelmää, kun siinä mainitaan nyt huomioitavina asioina kaikki objektiiviset tekijät. Vrt. kuitenkin Stevens 2015, s. 102–105, jossa kirjoittaja argumentoi, että TSA:n 26 johdantokappaleen mukaisen tunnistettavuuden määritelmän mukaan aikaisemmin henkilötietodirektiivin aikana anonyymeistä tiedoista saattoi tulla henkilötietoja. Ks. myös tarkemmin henkilötietodirektiivin 26 johdantokappaleen tulkinnasta Beyleveld – Townend 2004, jossa kirjoittajat jopa ehdottavat kyseisen johdantokappaleen toteamista henkilötietodirektiivin 1.1 artiklan vastaiseksi.

[324] C-582/14 Breyer, kohta 46.

[325] TSA 26 johdantokappale. Ks. myös Purtova 2018, s. 44.

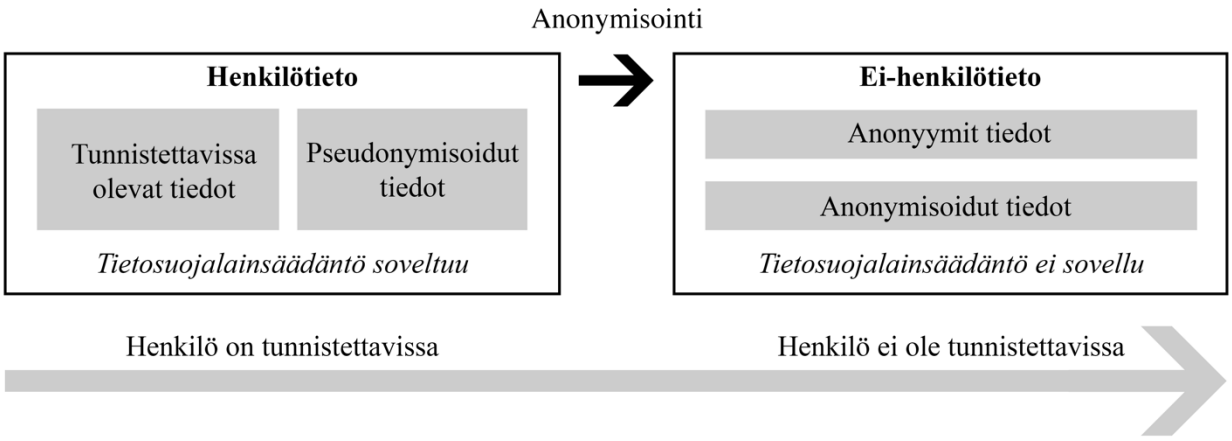
[326] WP 216, s. 6.

[327] WP 216. s. 6. Ks. myös Hintze 2018, s. 89.

[328] Tämän vahvistaa myös EUT ja julkisasiamies Sánchez-Bordona ratkaisussa C-582/14 Breyer.

[329] WP 216, s. 10. Tietosuojatyöryhmän tulkinnan voidaan tässä yhteydessä katsoa eroavan sen aiemmin WP 136:ssä omaksumasta tulkinnasta, jonka mukaan rekisterinpitäjän tietojenkäsittelyn tarkoituksella olisi merkitystä tunnistettavuuden arvioinnissa. Vrt. WP 136, s. 16. Toisaalta tämä voi myös korostaa tunnistettavuuden arvioinnin kontekstisidonnaisuutta, jota tietosuojatyöryhmäkin on useasti alleviivannut. Ks. myös Bolognini – Bistolfi 2017, s. 175.

Henkilötietojen anonymisoinnin määritelmästä käy ilmi, että sen arvioinnin painopiste on lopputuloksessa.[330] Henkilötiedoista tulee anonymisoinnin myötä olla poistettu kaikki luonnollisen henkilön tunnistamisen mahdollistavat tekijät, jolloin jäljelle jää vain anonyymejä tietoja. Anonyymien tietojen ja anonymisoitujen tietojen välinen ero on siinä, että anonyymit tiedot voivat olla joko olleet aina anonyymejä tietoja, jolloin ne eivät ole koskaan olleet tietosuojalainsäädännön piirissä, tai ne ovat voineet olla aikaisemmin henkilötietoja, mutta niistä on anonymisoinnin keinoin poistettu tunnistettavuus siten, etteivät ne ole enää henkilötietoja. Anonymisoidut tiedot ovat sen sijaan tietosuoja-asetuksen valossa olleet aina henkilötietoja ennen kuin niistä on poistettu tunnistettavuus anonymisoimalla. Tällaiset tiedot ovat aikaisemmin liittyneet tunnistettuihin tai tunnistettavissa oleviin henkilöihin ennen kuin rekisterinpitäjä tai kolmas on anonymisoinut kyseiset tiedot. Huomionarvoista on, että anonymisoitujen tietojen käsittelyyn liittynee useimmissa tapauksissa suurempi jäännösriski rekisteröidyille tunnistamisesta, kuin sellaisten anonyymien tietojen, jotka eivät ole koskaan olleet tietosuojalainsäädännön piirissä.[331] Henkilötietojen ja anonyymien tietojen välistä suhdetta kuvaa alla oleva kaavio.



Kaavio 1: Henkilötietojen ja anonyymien tietojen välinen suhde

Kuten kaaviosta käy ilmi, henkilötietojen ja anonyymien tietojen välistä suhdetta määrittelee perustavanlaatuisesti tunnistettavuus, joka käytännössä ratkaisee kussakin tilanteessa sen, ovatko tiedot

[330] WP 216, s. 5.
[331] Tämä johtuu tosin suurelta osin käytetystä anonymisointitekniikasta. Esimerkiksi aggregoidun tilastotiedon käsittelyn on katsottu olevan käytännössä riskitöntä. Sen sijaan esimerkiksi laajojen tietoaineistojen avoin julkaiseminen anonymisoituna sisältää käytännössä merkittäviä riskejä, kuten tutkielman 3.2.5 kappaleen esimerkit osoittavat.

henkilötietoja.[332] Kaaviossa henkilötiedot ovat kahdessa erillisessä kategoriassa sen havainnollistamiseksi, että henkilötiedoiksi tulkitaan myös sellaiset tiedot, joista henkilöitä ei ole mahdollista tunnistaa ilman lisätietoja, eli aiemmin kappaleessa 2.3 tarkastellut pseudonymisoidut tiedot. On kuitenkin tärkeää korostaa, että henkilötiedoille ei ole olemassa mitään erillisiä kategorioita, vaan tietosuojalainsäädännön näkökulmasta kaikenlaisiin henkilötietoihin sovelletaan samoja sääntöjä.[333]

Toisella puolella kaaviota ovat taas ei-henkilötiedot, eli sekä sellaiset tiedot, jotka eivät ole koskaan olleet tietosuojalainsäädännön piirissä, että tiedot, jotka on anonymisoitu tietosuojalainsäädännön vaatimukset täyttävällä tavalla. Näiden tietojen erityispiirteenä on, ettei niiden perusteella ole mahdollista tunnistaa henkilöitä, joten niiden käsittely ei vaikuta yhdenkään luonnollisen henkilön etuihin tai oikeuksiin. Tästä syystä kaavion oikealla puolella oleviin tietojen ei sovelleta henkilötietojen suojaa koskevia säännöksiä.

3.2 Henkilötietojen anonymisointi

3.2.1 Anonymisointi prosessina

Kuten edellä mainittiin, TSA:ssa anonymisoinnilla tarkoitetaan prosessia, jonka seurauksena henkilötiedoista poistetaan tai muutetaan riittävästi elementtejä siten, että tiedoista ei ole enää mahdollista tunnistaa luonnollisia henkilöitä.[334] Henkilötietojen anonymisoinnin arvioinnin painopiste on sen peruuttamattomassa lopputuloksessa ja niin tiedot anonymisoineen kuin kolmannen ulkopuolisen tulee olla kykenemätön tunnistamaan yksilöitä tiedoista tai palauttamaan tunnistettavuutta.[335] Tehokkaan ja riittävän anonymisoinnin kynnystä nostaa merkittävästi se, että yksilöiden tunnistamattomuuden tulee säilyä, vaikka tiedot yhdistettäisiin toisiin tietoihin, jotka ovat joko julkisia, taikka vain rekisterinpitäjän tai kolmannen hallussa.[336] Näin ollen henkilötietojen anonymisoinnista vastaavan tulee arvioida, millä keinoin ja millaisiin olemassa oleviin julkisiin tai ei-julkisiin tietoihin anonymisoidut tiedot voisi olla mahdollista *kohtuullisen todennäköisin keinoin* yhdistää siten, että yksilöitä olisi mahdollista tunnistaa tiedoista. Tämä henkilötietojen anonymisoinnille asetettu vaatimus tekee eurooppalaisen tietosuojalainsäädännön vaatimukset täyttävästä anonymisoinnista käytännössä erittäin

[332] WP 216, s. 5.
[333] Hintze 2018, s. 89.
[334] TSA 26 johdantokappale. Ks. myös TSA 4(1)(5) henkilötietojen pseudonymisoinnista prosessina.
[335] WP 216, s. 5. Ks. Esayas 2015, s. 3.
[336] WP 216, s. 3; C-582/14 Breyer, kohta 43. Ks. myös ICO 2012, s. 18, jossa tietojen yhdistelystä on todettu, että "*It is worth stressing that the risk of re-identification through data linkage is essentially unpredictable because it can never be assessed with certainty what data is already available or what data may be released in the future*".

vaikeaa, sillä rekisterinpitäjän on todella vaikea arvioida, millaisia muita tunnistamisen mahdollistavia tietoja voisi ylipäänsä olla olemassa.[337]

Edellä mainituista syistä henkilötietojen anonymisointi on prosessi, jonka toteuttaminen vaatii rekisterinpitäjältä tarkan arvioinnin tekemistä siihen liittyvistä riskeistä. Tämän prosessin tukena on tietosuojatyöryhmän anonymisointitekniikoita koskeva lausunto 5/2014 (WP 216), jonka mukaan tietosuojalainsäädännön vaatimukset täyttävän henkilötietojen anonymisoinnin yhteydessä tulee erityisesti huomioida neljä päätekijää:

> *”1) Anonymisointi voi olla tulosta henkilötietojen käsittelystä, jonka tavoitteena on peruuttamattomasti estää rekisteröidyn tunnistaminen,*
>
> *2) useat anonymisointitekniikat ovat mahdollisia, koska EU:n lainsäädännössä ei ole asiaa koskevia säännöksiä,*
>
> *3) kontekstuaalisiin elementteihin on tärkeää kiinnittää huomiota: On otettava huomioon ”kaikki” keinot, joita rekisterinpitäjä tai kolmannet osapuolet voivat ”kohtuullisesti” toteuttaa. On erityisesti kiinnitettävä huomiota siihen, mitä on nykytekniikan kehityksen valossa pidettävä ”kohtuullisesti toteutettavissa” olevana (ottaen huomioon tietokoneiden teho ja käytettävissä olevien välineiden lisääntyminen),*
>
> *4) anonymisointiin liittyy aina riskitekijä, joka on otettava huomioon kunkin anonymisointitekniikan pätevyyden arvioinnissa, ja myös tällaisen tekniikan avulla anonymisoitujen tietojen mahdolliset käytöt on otettava huomioon, sekä riskin vakavuus ja todennäköisyys arvioitava.”* [338]

Näiden neljän tietosuojatyöryhmän asettaman tekijän valossa anonymisointi on mahdollista nähdä prosessina, joka on henkilötietojen käsittelyä, jonka seurauksena rekisteröityjen tunnistaminen estyy peruuttamattomasti.[339] Tietosuojalainsäädäntö ei rajoita käytettävien anonymisointitekniikoiden joukkoa, vaan minkä tahansa tunnistettavuuden poistavan teknologian hyödyntäminen on mahdollista

[337] Ks. tähän liittyen Esayas 2015, s. 3.
[338] WP 216, s. 6–7.
[339] Ibid, s. 3.

henkilötietojen anonymisoinnin yhteydessä.[340] Anonymisointiprosessin arviointi perustuu kontekstuaalisiin elementteihin, erityisesti *kohtuullisen todennäköisiin keinoihin*[341], joiden arvioinnissa tulee tukeutua muun muassa asiaa koskeviin EUT:n vahvasti velvoittaviin ratkaisuihin. Näin ollen henkilötietojen anonymisointia tulee arvioida TSA:ssa yleisesti omaksutun riskiperusteisen lähestymistavan lähtökohdista käsin.[342] Tämä tarkoittaa muun muassa sen tarkastelemista, millaisen riskin anonymisoitujen tietojen käsittely aiheuttaa tiedoissa olevilla henkilöille.

Tietosuojatyöryhmän anonymisointitekniikoita koskevaa lausuntoa on kuitenkin myös kritisoitu siinä käytetystä terminologiasta, joka viittaisi siihen, että anonymisoinnin yhteydessä hyväksyttävä riski olisi tilanne, jossa ei olisi riskiä lainkaan.[343] Käytännössä lausunnossa ei oteta kantaa siihen, mikä on riittävä riskin taso, vaan lausunnossa periaatteessa vain esitellään erilaisia lähestymistapoja ja tekniikoita, joita hyödyntämällä anonymisointi voisi olla mahdollista saavuttaa. Tietosuojatyöryhmän käyttää tässä yhteydessä terminologiaa, joka viittaisi täydelliseen riskittömyyteen, kuten ”*irreversibly prevent identification*”[344] ja ”*in order to assess whether the anonymisation process is sufficiently robust, i.e. wheter identification has become 'reasonably' impossible*”.[345] Tällainen täydellinen riskittömyys anonymisoinnin arvioinnissa olisi kuitenkin ristiriidassa TSA:ssa yleisesti omaksutun riskiperusteisen lähestymistavan, TSA:n 26 johdantokappaleen ja EUT:n Breyer-ratkaisun kanssa. Lisäksi tietosuojatyöryhmä itsekin toteaa lausunnossaan, että anonymisoinnin luonteeseen kuuluu riskitekijän olemassaolo.[346]

Tehokkaan henkilötietojen anonymisoinnin yhteydessä alkuperäiset tiedot, joista on edelleen mahdollista tunnistaa henkilöitä, on poistettava. Jos rekisterinpitäjä säilyttää alkuperäiset tiedot, ovat tiedot lähtökohtaisesti edelleen tietosuojalainsäädännön näkökulmasta henkilötietoja, sillä vähintään rekisterinpitäjä pystyy tunnistamaan tiedoista henkilöitä, jolloin tiedot ovat pikemminkin pseudonymisoituja henkilötietoja. Toisaalta mielenkiintoinen kysymys edellä 2.2.4 kappaleessa käsitellyn Breyer-ratkaisun valossa on, voisiko tietoja olla mahdollista käsitellä saman organisaation sisällä

[340] Käsittelen yleisimpiä anonymisointitekniikoita seikkaperäisesti tutkielman kappaleessa 3.2.4.
[341] Kuten edellä on mainittu, eurooppalaisessa tietosuojalainsäädännössä anonymisoinnin tehokkuutta voidaan katsoa arvioitavan riskiperusteista lähestymistavasta käsin. Ks. tähän liittyen esim. Stalla-Bourdillon – Knight 2016, s. 286.
[342] TSA:ssa omaksuttu riskiperusteisuus ilmenee muun muassa asetuksen 24, 25 ja 35 artikloissa. Tarkemmin TSA:n riskiperusteisesta lähestymistavasta ks. esim. Quelle 2018, s. 37–42.
[343] Ks. esim. El Emam – Álvarez 2015, s. 74–76; Esayas 2015, s. 6.
[344] WP 216, s. 3.
[345] Ibid, s. 8. Ks. El Emam – Álvarez 2015, s. 75, jossa kirjoittajat toteavat, että käsitteitä ”*reasonableness*” ja ”*impossible*” on todella vaikea edes tulkita yhdessä.
[346] WP 216, s. 7.

sekä tunnisteellisina että tunnistamattomina, jos organisaation sisällä tunnisteellisten tietojen käsittelyä rajoitettaisiin vain tietyille osastoille.[347]

Esimerkiksi jos tuotekehitystä tekevässä yrityksessä tai tieteellistä tutkimusta tekevässä sairaalassa tietojen käsittely olisi toteutettu siten, että vain tuotekehityksestä tai tutkimuksesta vastaavalla henkilöllä olisi pääsy tunnisteellisiin henkilötietoihin, ja muut varsinaisen tutkimuksen tekijät käsittelisivät ainoastaan anonymisoituja/pseudonymisoituja tietoja, voisiko työntekijöiden suorittama käsittely olla TSA:n soveltamisalan ulkopuolella.[348] Tällaisessa tapauksessa henkilötietojen käsittely tapahtuisi saman rekisterinpitäjän sisällä, jolloin tietosuojalainsäädännön näkökulmasta tiedot olisivat lähtökohtaisesti pseudonymisoituja henkilötietoja. Jos toisaalta tunnistamattomassa muodossa olevilla työntekijöillä ei ole tosiasiallisesti minkäänlaisia keinoja tunnistaa luonnollisia henkilöitä, sillä vastaava henkilö olisi esimerkiksi salassapitovelvollinen, voisi tällaisia tietoja olla teoriassa mahdollista käsitellä anonyymeinä tietoina Breyer-ratkaisun valossa.[349]

Joissain tilanteissa pitkälle viety tunnistettavuuden tulkitseminen ainoastaan yhden rekisterinpitäjän näkökulmasta voi myös johtaa erikoisiin lopputuloksiin. Muun muassa sellaisten organisaatioiden, joilla on toimintaa useilla eri toimialoilla, tulkitseminen yhdeksi rekisterinpitäjäksi voi johtaa tilanteeseen, jossa organisaatiossa käsitellään tietosuojaoikeudellisesta näkökulmasta laaja-alaisemmin tunnisteellisia tietoja, kuin organisaatiossa ollaan tietoisia. Jos esimerkiksi suuren kaupungin sosiaalitoiminnoissa käsitellään tietoja, joista henkilöitä on mahdollista tunnistaa, olisivat vastaavat tiedot kaupungin elinkeinotoiminnassa henkilötietoja, vaikka elinkeinotoiminnassa ei oltaisi edes tietoisia sosiaalitoiminnoissa käsitellyistä tiedoista. Näin ollen pelkästään yhtenä oikeushenkilönä toimivien

[347] Kysymys liittyy EUT:n Breyer-ratkaisussa omaksumaan tulkintaan, että samat tiedot voivat olla samaan aikaan toiselle taholle henkilötietoja ja toiselle ei, sillä jälkimmäinen ei kykene tunnistamaan henkilöitä tiedoista. Ks. myös Mourby et al. 2018, s. 233, jossa kirjoittajat toteavat, että Breyer-ratkaisu jättää avoimeksi mahdollisuuden käsitellä henkilötietoja pseudonymisoituina yhden organisaation sisällä ja luovuttaa tietoja toiselle organisaatiolle siten, että tiedot olisivat anonyymejä vastaanottavan organisaation näkökulmasta.

[348] Tämänkaltaisista organisaation sisäisistä tietojen liikkumisen rajoituksista käytetään yleisesti termiä ”*Chinese walls*”, joilla muun muassa investointipankeissa kontrolloidaan eturistiriitojen syntymistä eri osastojen välillä. Ei ole kuitenkaan selkeää, että toimisiko samantapainen konsepti tietosuojalainsäädännön kontekstissa.

[349] Periaatteessa samaan tulkintaan voi päätyä tietosuojatyöryhmän henkilötiedon käsitettä koskevaa lausuntoa tulkitsemalla. Ks. WP 136, s. 15–16, jossa tietosuojatyöryhmä käyttää esimerkkinä potilaiden terveystietojen siirtämistä sairaalalta yksityiselle yritykselle lääketieteellistä tutkimusta varten. Tietosuojatyöryhmä esittelee tilanteen, jossa potilaita ei ole tiedoissa yksilöity nimellä, vaan satunnaisilla numeroilla, että tietojen yhtenäisyys säilyy ja potilaat eivät sekoitu toisiinsa. Lisäksi potilaiden nimet säilyvät ainoastaan vastaavien salassapitovelvollisten lääkärien tiedossa, tietoihin ei sisälly tunnistamisen mahdollistavia lisätietoja ja tietojen käsittelyssä on toteutettu asianmukaiset oikeudelliset, tekniset ja organisatoriset toimet tunnistettavuuden estämiseksi. Tällaisessa tilanteessa tietosuojatyöryhmän mukaan voi olla mahdollista, että tiedot vastaanottaneella yrityksellä ei ole käytössään kohtuullisen todennäköisiä keinoja rekisteröityjen tunnistamiseksi, jolloin kyseisiä tietoja voisi käsitellä anonyymeinä tietoina. Käytännössä ainut ero tällaiseen yrityksen suorittamaan henkilötietojen käsittelyyn verrattuna organisaation sisällä rajattuun käsittelyyn on, että organisaation sisällä tietoja käsitellään saman rekisterinpitäjän toimesta.

rekisterinpitäjien näkökulmasta tunnistettavuuden arvioiminen vain organisaation oikeudellisen muodon perusteella voi johtaa epätarkoituksenmukaisiin tulkintoihin tunnistettavuuden näkökulmasta – etenkin, jos rekisterinpitäjän toiminta on laaja-alaista.[350]

Yhdistyneen Kuningaskunnan tietosuojaviranomainen (Information Comissioner's Office, ICO) on anonymisointia käsittelevässä vuonna 2012 julkaistussa oppaassaan kehittänyt anonymisoinnin arvioinnissa käytettävän testin "*The motivated intruder test*".[351] Tämä tarkoittaa käytännössä anonymisoinnin tehokkuuden arvioimista siitä tilanteesta käsin, että motivoitunut hyökkääjä pyrkisi tunnistamaan anonymisoiduista tiedoista luonnollisia henkilöitä. Testin lähtökohtana on, että motivoitunut hyökkääjä on henkilö, jolla ei etukäteistä tietoa, mutta joka haluaa tunnistaa henkilöitä anonymisoiduista tietoista. Motivoitunut hyökkääjä olisi kohtuullisen pätevä ja hänellä olisi käytössään normaalit resurssit, kuten pääsy internettiin, kirjastoihin ja kaikkiin julkisiin asiakirjoihin, jonka lisäksi hän voisi suorittaa tutkintaa tietoihin liittyen. Hyökkääjä voisi esimerkiksi lähestyä sellaisia henkilöitä, joilla voisi olla lisätietoja anonymisoiduista tiedoista ja houkutella lisätietoja tietäviä henkilöitä kertomaan tietonsa. Motivoituneella hyökkääjällä ei olisi kuitenkaan erityisosaamista, kuten kompetenssia tehdä tietomurtoja, pääsyä erikoisvälineisiin, eikä hän syyllistyisi rikoksiin, kuten murtautumiseen tiloihin, jossa tietoja säilytetään turvallisesti.[352]

TSA:n sääntelyä vastaavasti motivoitunutta hyökkääjä koskevassa testissä tietosuojaa arvioidaan riskiperusteisesti.[353] Tästä johtuen mitä arkaluontoisemmista tiedoista on kyse, sitä suurempi riski on, että motivoitunut hyökkääjä pyrkii tunnistamaan henkilöitä anonymisoiduista tiedoista. Tällaisia arkaluontoisia tietoja olisivat lähtökohtaisesti ainakin TSA 9 artiklan mukaiset erityiset henkilötietoryhmät,[354] kuten esimerkiksi tieto siitä, että henkilöllä on jokin tarttuva tauti. Arkaluontoisten tietojen yhteydessä motivoitunut hyökkääjä pyrkii myös todennäköisemmin käyttämään pidemmälle meneviä

[350] Rekisterinpitäjyyden määräytymistä organisaation oikeushenkilöllisyyden perusteella on oikeuskirjallisuudessa toisaalta kannatettu, toisaalta kritisoitu. Lähtökohtaisesti vaikuttaa siltä, että rekisterinpitäjyyden määräytymisen arvioinnissa on käytännöllisempää huomioida muutkin seikat kuin pelkästään organisaation oikeushenkilöllisyys. Tämä olisi perusteltua myös tunnistettavuuden tulkinnan näkökulmasta. Ks. tähän liittyvästä keskustelusta Van Alsenoy 2019, s. 121–124.
[351] ICO 2012, s. 22.
[352] Ibid, s. 22–23.
[353] Ibid, s. 21.
[354] TSA 9 artiklassa säädetään erityisiä henkilötietoryhmiä koskevasta käsittelystä, jonka 1 kohdan mukaan "*Sellaisten henkilötietojen käsittely, joista ilmenee rotu tai etninen alkuperä, poliittisia mielipiteitä, uskonnollinen tai filosofinen vakaumus tai ammattiliiton jäsenyys sekä geneettisten tai biometristen tietojen käsittely henkilön yksiselitteistä tunnistamista varten tai terveyttä koskevien tietojen taikka luonnollisen henkilön seksuaalista käyttäytymistä ja suuntautumista koskevien tietojen käsittely on kiellettyä."* Kyseisen artiklan 2 kohdassa säädetään sen sijaan tilanteista, jolloin 1 kohtaa ei sovelleta, eli käsittely on sallittua.

keinoja, kuin tavanomaisten henkilötietojen yhteydessä. Arkaluontoisten terveystietojen anonymisointi tulisi tästä syystä suorittaa tavallista vahvemmilla anonymisointitekniikoilla.[355] Motivoituneen hyökkääjän käyttäminen anonymisoinnin tehokkuuden arvioinnissa on hyödyllinen apuväline, sillä ICO on sen kehittämisen yhteydessä huomioinut silloisen tietosuojalainsäädännön vaatimukset, kuten henkilötietodirektiivin 26 johdantokappaleen, joka vastaa relevanteilta osin TSA:n 26 johdantokappaleen sääntelyä. Näin ollen motivoitunutta hyökkääjää koskevan testin voidaan katsoa vastaan Ison-Britannian tietosuojaviranomaisen näkemystä siitä, mitä *kohtuullisen todennäköisillä keinoilla* tarkoitetaan.[356]

3.2.2 Anonymisoinnin hyödyt

3.2.2.1 Henkilötietojen anonymisoinnin hyödyistä yleisesti

Henkilötietojen anonymisoinnin tarkastelun yhteydessä on tärkeää ymmärtää, miksi erilaiset toimijat haluavat anonymisoida hallussaan olevia henkilötietoja. Tässä yhteydessä huomionarvoista on, että niin anonymisoinnista seuraavat hyödyt kuin tavoitteet, joihin anonymisoinnilla pyritään, ovat verrattain yhtenäisiä yksityisten ja julkisten toimijoiden välillä, sillä molemmat pyrkivät muun muassa tietojen hyödyntämiseen uusiin käyttötarkoituksiin tietosuojalainsäädännön estämättä.[357] Yksi tällainen uusi käyttötarkoitus on tietojen analysoiminen tilastotiedettä ja tietotekniikkaa hyödyntämällä.

Varsinkin julkisella sektorilla käsitellään tavanomaisesti huomattavia määriä erilaisia henkilötietoja erilaisten palvelujen järjestämisen yhteydessä. Tällaisten henkilötietojen anonymisoiminen mahdollistaisi sen, että kyseisiä tietoja voitaisiin analysoida laillisesti ja ihmisten luottamusta vaarantamatta.[358] Tämän seurauksena tietojen perusteella voisi olla mahdollista löytää ennen tuntemattomia korrelaatioita esimerkiksi sen ennustamiseksi, kuinka todennäköisesti sama henkilö palaa myöhemmin sosiaalihuollon asiakkaaksi.[359] Lisäksi tiedoista saattaisi olla pääteltävissä ja löydettävissä ne

[355] Anonymisoitujen arkaluontoisten tietojen julkaisemiseen liittyvässä päätöksenteossa tulisi myös olla erityisen varovainen ja arvioida julkaisemiseen liittyvät riskit tarkasti. Ks. tähän liittyen Ruotsin Puolustustutkimuslaitoksen (Totalförsvarets forskningsinstitut, FOI) tutkija Magnus Jändelin artikkeli ”Decision support for releasing anonymised data” vuodelta 2014.

[356] Lisäksi EUT oli ainakin ICO:n motivoitunutta hyökkääjää koskevan testin kanssa samaa mieltä siitä, että kohtuullisen todennäköisillä keinoilla ei tarkoiteta keinoja, jotka ovat erikseen laissa kiellettyjä. Ks. C-582/14 Breyer, kohta 46.

[357] Ks. esim. TfL heinäkuu 2019.

[358] Ks. UKAN 2016, s. 7, jossa anonymisointia lähestytään siitä näkökulmasta, että sen avulla on mahdollista ylläpitää luonnollisten henkilöiden luottamusta heidän tietojenkäsittelyään kohtaan. Ks. myös Euroopan komissio 2018, s. 3, jossa komissio lausuu, että tehokkaalla tietosuojasääntelyllä pyritään vahvistamaan yksilöiden luottamusta yhteiskuntaa kohtaan.

[359] HS 7.6.2018. Esimerkiksi Espoon kaupunki ja ohjelmistotalo Tieto tekivät vuonna 2018 tekoälykokeilun, jonka seurauksena löydettiin 280 tekijää, jotka ennakoivat tulevaa lastensuojelupalvelujen tarvetta. Kokeilu perustui julkishallinnon asiakkuusdatan analysoimiseen ja aineistoon sisältyi kaikki Espoon kaupungin hallussa oleva sosiaali- ja terveystieto

muuttujat, jotka viittaavat suureen riskiin tulevasta sosiaalipalvelujen tarpeesta. Näin ollen henkilötietojen anonymisointi voisi mahdollistaa henkilötietojen analysoimisen siten, että potentiaaliset avun tarvitsijat voitaisiin saattaa julkisten palvelujen tarjoaman avun piiriin ennaltaehkäisevästi.[360] Toiseksi julkisen sektorin toimijat pyrkivät mahdollisuuksien mukaan julkaisemaan hallussaan olevia tietovarantoja myös avoimena datana[361], sillä avoin data edistää muun muassa sekä uudenlaisten innovaatioiden syntymistä että hallinnon avoimuutta.[362] Tämä edellyttää usein julkaistavissa tiedoissa olevien henkilötietojen anonymisointia, jotta tietojen julkaisu ei loukkaa luonnollisten henkilöiden yksityisyyden suojaan ja henkilötietojen suojaan liittyviä oikeuksia.[363]

Henkilötietojen anonymisointi voi olla erittäin tavoiteltavaa myös yksityisen sektorin toimijoille. Nykyaikana useat yritykset keräävät esimerkiksi asiakkaistaan huomattavan määrän erilaisia tietoja, joista merkittävä osa on tietosuojalainsäädännön näkökulmasta henkilötietoja.[364] Tietojen anonymisoinnin myötä yritykset voisivat analysoida muun muassa toimintansa yhteydessä kerättyjä paikkatietoja, terveystietoja, asiakastietoja tai kuvia käytännössä rajoituksessa.[365] Tämän seurauksena yritysten olisi mahdollista tehostaa toimintaansa tai havaita täysin uusia ansaintakeinoja esimerkiksi asiakkaiden matkustuskäyttäytymisen analysoinnin perusteella.[366] Näin ollen tietosuojalainsäädännön vaatimukset täyttävä anonymisointi voi tuottaa monenlaisia hyötyjä useille eri toimijoille, jotka käsittelevät henkilötietoja toimintansa yhteydessä ja myös tietosuojatyöryhmä on myöntänyt anonymisoinnin potentiaalisen arvon yrityksille ja yhteiskunnalle tietojen hyödyntämisen strategiana.[367]

---

sekä täydentävä asiakkuusdata vuosien 2002 ja 2016 väliltä. Aineistossa oli tietoa yli 500 tuhannesta henkilöstä sekä yli 37 miljoonasta yksittäisestä asiakaskontaktista. Tieto Oyj:n asiasta julkaisen tiedotteen perusteella ”Henkilöitä yksilöivät tiedot, kuten nimi, henkilötunnus ja osoite, salattiin datasta jo tiedonhakuvaiheessa”. Vaikuttaa siltä, että tietoja käsiteltiin anonymisoituina tietoina, sillä niistä poistettiin suorat tunnisteet. Kokeilussa käytetyt tiedot on sittemmin tuhottu, joten jää selvittämättä, olivatko analysoidut tiedot todellisuudessa anonymisoitu tietosuojalainsäädännön vaatimukset täyttävällä tavalla.

[360] Tähän liittyen Suomi on toteuttamassa kansallista tekoälyohjelma AuroraAI:ta, ks. Valtiovarainministeriö 2019.

[361] Zuiderveen Borgesius et al. 2015, s. 2088–2091. Ks. myös Oswald 2014, s. 246–247.

[362] Mäenpää 2016, s. 10–11. Ks. tähän liittyen julkisen sektorin hallussa olevien tietojen uudelleenkäytöstä annettu direktiivi 2003/98/EY (PSI-direktiivi), jonka uudistamistyö on parhaillaan käynnissä EU:ssa.

[363] Perusoikeuskirja 7 ja 8 artiklat; EIS 8 artikla.

[364] Kriittiset kirjoittajat oikeuskirjallisuudessa ovat jopa argumentoineet, että lähitulevaisuudessa kaikki tieto tulee olemaan henkilötietoa. Ks. Purtova 2018, jossa kirjoittaja argumentoi tietosuojalainsäädännöstä tulevan ”*The law of everything*”. Tästäkin näkökulmasta tietojen anonymisointi tulee olemaan elintärkeää.

[365] Tässä yhteydessä on tosin huomioitava myös sektorikohtainen erityislainsäädäntö, joka voi aiheuttaa lisävaatimuksia käsittelylle. Esimerkiksi terveystietojen käsittelyyn liittyy huomattava määrä huomioon otettavaa erityislainsäädäntöä.

[366] Esimerkiksi joukkoliikenteen palveluja tarjoava yritys voisi anonymisoituja paikkatietoja analysoimalla havaita matkustajien liikkumisessa uusia johdonmukaisuuksia, joiden perusteella se voisi optimoida palvelujaan lisäämällä liikennöintiä tietyllä reittivälillä.

[367] WP 216, s. 3.

Kuten jo edellä todettiin, anonyymit tiedot eivät kuulu tietosuojalainsäädännön soveltamisalan piiriin, joten tällaisia tietoja voi käsitellä ilman henkilötietojen suojaa koskevan lainsäädännön velvoitteita.[368] Tämä pätee yhtä lailla anonymisoituihin tietoihin, joten henkilötietojen anonymisoinnilla on mahdollista käytännössä rajata koko tietosuoja-asetuksen soveltaminen tietojen käsittelyn ulkopuolelle.[369] On mahdollista argumentoida, että merkittävimpiä henkilötietojen anonymisoinnista seuraavia hyötyjä TSA:n näkökulmasta ovat 5 artiklan mukaisten henkilötietojen käsittelyä koskevien periaatteiden[370] ja TSA 6 artiklan mukaisten henkilötietojen käsittelyperusteiden[371] vaatimusten poistuminen henkilötietojen käsittelystä. Tämä johtuu siitä, että kyseiset artiklat vaikuttavat etenkin tietojen keräämiseen, hyödyntämiseen ja säilyttämiseen. Tarkastelen anonymisoinnin hyötyjä juuri näiden kyseisten artiklojen kannalta, sillä niiden on mahdollista katsoa vaikuttavan TSA:n artikloista eniten sellaiseen henkilötietojen käsittelyyn, joka on anonymisoinnin näkökulmasta relevanttia.

3.2.2.2 Anonymisoinnin hyödyt tietosuoja-asetuksen 5 ja 6 artiklojen näkökulmasta

Yksi henkilötietojen käsittelyä koskevista merkittävimmistä periaatteista on *käyttötarkoitussidonnaisuuden periaate*, josta säädetään TSA:n 5(1)(b) artiklassa.[372] Käyttötarkoitussidonnaisuudella tarkoitetaan sitä, että henkilötiedot on kerättävä tiettyä nimenomaista ja laillista tarkoitusta varten, eikä niitä saa käsitellä myöhemmin näiden tarkoitusten kanssa yhteensopimattomalla tavalla.[373] TSA:n mukainen käyttötarkoitussidonnaisuus muodostuu kahdesta rekisterinpitäjälle asetetusta vaatimuksesta: 1) henkilötietoja saa *kerätä* vain tiettyä nimenomaista ja laillista tarkoitusta varten, jonka lisäksi 2) henkilötietoja ei saa *käsitellä* alkuperäisen käyttötarkoituksena kanssa yhteensopimattomalla tavalla.[374]

---

[368] Tietosuoja-asetus 26 johdantokappale.

[369] On kuitenkin huomattava, että anonyymeihinkin tietoihin voi tulla sovellettavaksi esimerkiksi sellainen sähköisen viestinnän tietosuojaa koskeva lainsäädäntö, jota sovelletaan kaikkiin tietoihin riippumatta siitä, ovatko kyseiset tiedot henkilötietoja vai ei.

[370] Tarkastelen henkilötietojen anonymisoinnista rekisterinpitäjälle seuraavia hyötyjä kolmen merkittävän tietosuojaperiaatteen näkökulmasta, jotka ovat 1) *käyttötarkoitussidonnaisuus*, 2) *säilytyksen rajoittaminen*, ja 3) *eheys ja luottamuksellisuus*.

[371] TSA 6 artiklassa säädetään käsittelyn lainmukaisuuden edellytyksistä, joista käytännössä yleisimpiä, eli 6(1)(a) artiklan mukaista *suostumusta* ja 6(1)(f) artiklan mukaista *oikeutettua etua* käsittelen tutkielmassa henkilötietojen anonymisoinnin näkökulmasta.

[372] Purtova 2014, s. 14.

[373] TSA 5(1)(b). Kohdassa säädetään myös poikkeuksesta yhteensopimattomuuteen liittyen, jonka mukaan *"myöhempää käsittelyä yleisen edun mukaisia arkistointitarkoituksia taikka tieteellisiä tai historiallisia tutkimustarkoituksia tai tilastollisia tarkoituksia varten ei katsotan 89 artiklan 1 kohdan mukaisesti yhteensopimattomaksi alkuperäisten käyttötarkoitusten kanssa"*.

[374] WP 203, s. 3. Ks. myös Mayer-Schönberger – Padova 2016, s. 325. Käyttötarkoitussidonnaisuuden periaatteesta henkilötietodirektiivin sääntelyssä ks. Pitkänen et al. 2014, s. 79–80.

Ensimmäisen käyttötarkoitussidonnaisuuden periaatteen asettama vaatimus rajoittaa henkilötietojen keräämisen vain tiettyyn etukäteen tiedossa olevaan lailliseen tarkoitukseen, jonka puitteissa henkilötietojen käsittelyn tulee tapahtua, kun toinen vaatimus taas kohdistuu henkilötietojen myöhempään käyttöön.[375] Arvioitaessa henkilötietojen myöhemmän käytön yhteensopivuutta alkuperäisen käyttötarkoituksen kanssa, tulee tietosuojatyöryhmän mukaan ottaa huomioon neljä keskeistä kriteeriä: ”*1) uuden käyttötarkoituksen suhde alkuperäiseen lainmukaiseen, 2) henkilötietojen alkuperäisen keräämisen tilanne ja rekisteröidyn kohtuulliset odotukset tietojen käyttötarkoituksesta, 3) tietojen luonne ja tietojen käsittelyn vaikutus rekisteröityyn sekä 4) rekisterinpitäjän keinot reilun ja lainmukaisen käsittelyn turvaamiseksi ja käsittelyn liiallisten vaikutusten estämiseksi.”* [376] Näin ollen henkilötietojen myöhempää hyödyntämistä arvioitaessa on erityisesti huomioitava muun muassa tietojen arkaluontoisuus, sillä rekisteröity mitä todennäköisimmin olettaa, ettei arkaluontoisia tietoja käytetä muuhun kuin siihen nimenomaiseen käyttötarkoitukseen, jota varten kyseiset tiedot on kerätty.

Käyttötarkoitussidonnaisuuden ja henkilötietotietojen anonymisoinnin yhteydessä on huomioitava, että anonymisoinnin kohteena olevat henkilötiedot tulee olla joka tapauksessa alun perin kerätty tiettyä nimenomaista ja laillista tarkoitusta varten.[377] Jos tämä käyttötarkoitussidonnaisuuden ensimmäinen vaatimus ei täyty, on kyseisten henkilötietojen käsittely jo itsessään tietosuojalainsäädännön vastaista ja tietojen anonymisointi ei tee alkuperäisestä laittomasta keräämisestä sallittua. Tästä huolimatta sinänsä laittomasti kerättyjenkin henkilötietojen anonymisointi vähentää rekisteröidyille aiheutuneita riskejä, sillä anonymisoinnin seurauksena laiton henkilötietojen käsittely päättyy, kun tiedot eivät ole enää henkilötietoja.[378]

Käyttötarkoitussidonnaisuuden periaatteen toinen vaatimus on tärkeä periaate tehokkaan henkilötietojen suojan näkökulmasta, mutta se samalla rajoittaa kerättyjen tietojen myöhempää hyödyntämistä merkittävästi.[379] Tämän vaatimuksen kiertäminen on yksi tärkeimmistä tavoitteista, johon anonymisoinnilla pyritään, sillä aikaisemmin kerättyjen tietojen analysoiminen uusilla menetelmillä voi johtaa esimerkiksi uusiin havaintoihin ihmisten käyttäytymisessä.[380] Tietosuojalainsäädännön vaatimukset täyttävän henkilötietojen anonymisoinnin myötä rekisterinpitäjän ei tarvitse enää huomioida

[375] Purtova 2014, s. 14
[376] WP 203, s. 3; Korpisaari et al. 2018, s. 93.
[377] WP 216, s. 7.
[378] Tästä näkökulmasta voisi olla perusteltua anonymisoida sellaiset henkilötiedot, jotka on kerätty ennen TSA:n voimaantuloa aikanaan laillisesti, mutta joiden keräämisen käsittelyperuste ei enää nykyisin täyttäisi tietosuojalainsäädännön edellytyksiä ja esimerkiksi suostumuksen kerääminen rekisteröidyiltä erikseen ei olisi mahdollista.
[379] Mayer-Schönberger – Cukier 2013, s. 153–154. Ks. myös Mayer-Schönberger – Padova 2016, s. 330.
[380] Mayer-Schönberger – Cukier 2013, s. 53–54.

käyttötarkoitussidonnaisuutta anonymisoitujen tietojen käsittelyssä, jolloin rekisterinpitäjä voi käyttää tietoja myös mihin tahansa muuhun tarkoitukseen, kuin alkuperäiseen käyttötarkoitukseen, eikä rekisterinpitäjän tarvitse enää arvioida käsittelyn yhteensopivuutta alkuperäisen käyttötarkoituksen kanssa.

Toinen henkilötietojen anonymisoinnin näkökulmasta merkittävä tietosuojaperiaate on *säilytyksen rajoittaminen*, josta säädetään TSA 5(1)(e) artiklassa.[381] Säilytyksen rajoittaminen tarkoittaa sitä, että henkilötiedot on säilytettävä muodossa, josta rekisteröity on tunnistettavissa ainoastaan niin kauan kuin on tarpeen tietojenkäsittelyn tarkoituksen toteuttamista varten.[382] Henkilötietojen anonymisointi mahdollistaa sen, että rekisterinpitäjä voi periaatteessa säilyttää anonymisoituja tietoja rajoittamattoman ajan, sillä tällaiset tiedot tulkitaan tietosuojalainsäädännön näkökulmasta poistetuiksi. Lisäksi huomionarvoista on, että henkilötietojen anonymisoinnilla voi olla myös mahdollista toteuttaa rekisterinpitäjän velvollisuus, joka johtuu rekisteröidyn TSA 17(1) artiklan mukaisesta oikeudesta tietojen poistamiseen (”oikeus tulla unohdetuksi”).[383]

Tämän vahvisti Itävallan tietosuojaviranomainen (Die Datenschutzbehörde, DSB) 5. joulukuuta 2018 antamassaan ratkaisussa DSB-D123.270/0009-DSB/2018, jossa oli kyse siitä, onko henkilötietojen anonymisointi riittävä toimenpide TSA:n 17(1) artiklan mukaisen rekisteröidyn tietojen poistamista koskevan oikeuden toteuttamiseksi.[384] Tapauksessa Itävallan tietosuojaviranomainen päätyi tulkitsemaan TSA:n 4(1)(2) artiklan mukaisen henkilötietojen käsittelyn määritelmässä mainittua tietojen *tuhoamista* yhdessä TSA 17(1) artiklan kanssa siten, että henkilötietojen anonymisointi voi periaatteessa olla yksi keino henkilötietojen *poistamiseksi* TSA 17(1) artiklan nojalla.[385] Näin ollen DSB

[381] Säilytyksen rajoittaminen liittyy oleellisesti myös toiseen merkittävään tietosuojaperiaatteeseen: tietojen minimointiin. TSA 5(1)(c) mukaan ”*henkilötietojen on oltava asianmukaisia ja olennaisia ja rajoitettuja siihen, mikä on tarpeellista suhteessa niihin tarkoituksiin, joita varten niitä käsitellään (”tietojen minimointi”).*

[382] TSA 5(1)(e). Säilytyksen rajoittamista koskevassa kohdassa on vastaavanlainen yleisen edun mukaisia arkistointitarkoituksia taikka tieteellisiä tai historiallisia tutkimustarkoituksia tai tilastollisia tarkoituksia koskeva poikkeus kuin käyttötarkoitussidonnaisuuden periaatetta koskevassa b-alakohdassa. Kuitenkin säilytyksen rajoittamisen yhteydessä edellytetään myös sitä, että TSA:ssa vaaditut tekniset ja organisatoriset toimenpiteet on pantu täytäntöön rekisteröityjen oikeuksien ja vapauksien turvaamiseksi. Ks. TSA 25 artikla sisäänrakennetusta ja oletusarvoista tietosuojasta.

[383] TSA 17(1) artiklan mukaan ”*rekisteröidyllä on oikeus saada rekisterinpitäjä poistamaan rekisteröityä koskevat henkilötiedot ilman aiheetonta viivytystä, ja rekisterinpitäjällä on velvollisuus poistaa henkilötiedot ilman aiheetonta viivytystä, edellyttäen että jokin seuraavista perusteista täyttyy: a) henkilötietoja ei enää tarvita niihin tarkoituksiin, joita varten ne kerättiin tai joita varten niitä muutoin käsiteltiin; b) rekisteröity peruuttaa suostumuksen, johon käsittely on perustunut 6 artiklan 1 kohdan a alakohdan tai 9 artiklan 2 kohdan a alakohdan mukaisesti, eikä käsittelyyn ole muuta laillista perustetta;[…]”.*

[384] DSB 2018.

[385] Kyseisessä tapauksessa DSB:n tulkinnan mukaan yritys oli osittain tuhonnut ja osittain poistanut rekisteröityä koskevat henkilötiedot. DSB tarkoittaa tuhoamisella sitä, ettei yritys ollut säilyttänyt joitain tietoja rekisteröidyistä edes anonyymeinä tietoina, kun taas poistaminen tarkoittaa, että henkilö ei ole enää tunnistettavissa kyseisistä tiedoista. Ks. myös Data Protection Handbook 2018, s. 93.

katsoi, että kyseisessä tapauksessa yrityksen suorittama henkilötietojen anonymisointi täytti tietosuojalainsäädännön vaatimukset, kun henkilötiedot oli osittain ”tuhottu” ja osittain ”poistettu” siten, että poistetut tiedot korvattiin synteettisillä tiedoilla.[386]

Mielenkiintoista on, että DSB otti ratkaisussaan kantaa TSA 26 johdantokappaleessa mainittuun tunnistamisen yhteydessä huomioon otettavaan tekniseen kehitykseen kohtuullisen todennäköisten keinojen arvioinnissa. DSB:n näkemyksen mukaan hypoteettinen kehitys tunnistamisen mahdollistavassa teknologiassa ei vaikuta ratkaisun lopputuloksen arviointiin, eikä siten tee henkilötietojen poistamisesta riittämätöntä.[387] Tämä on TSA:n näkökulmasta lautunut tulkinta, sillä 26 johdantokappaleessa mainitaan eksplisiittisesti yhtenä tunnistettavuuden arvioinnissa huomioitavana keinona tekninen kehitys. Toisaalta DSB:n tulkinta on perusteltua tässä yksittäisessä tapauksessa, sillä tapauksessa kyseessä oleva yritys säilytti anonymisoituja tietoja vain yhdeksän kuukauden ajan, joten mitä todennäköisimmin tässä ajassa tunnistamisen mahdollistavissa teknologioissa ei tapahdu merkittävää kehitystä.

Kolmas anonymisoinnin kannalta merkittävä tietosuojaperiaate on *henkilötietojen eheys ja luottamuksellisuus*, joista säädetään TSA:n 5(1)(f) artiklassa. Henkilötietojen eheys ja luottamuksellisuus tarkoittavat sitä, että henkilötietoja on käsiteltävä tavalla, jolla varmistetaan henkilötietojen asianmukainen turvallisuus, mukaan lukien suojaaminen luvattomalta ja lainvastaiselta käsittelyltä sekä vahingossa tapahtuvalta häviämiseltä, tuhoutumiselta tai vahingoittumiselta käyttäen asianmukaisia organisatorisia tai teknisiä toimia.[388] Tämä periaate on henkilötietojen anonymisoinnin näkökulmasta relevantti, sillä vaikka anonymisointi ei täyttäisi tietosuojalainsäädännön vaatimuksia, jolloin kyseiset tiedot olisivat edelleen henkilötietoja, parantaisi anonymisointi käsiteltävien tietojen tietoturvaa

[386] Vaikuttaa siltä, että Suomessa tällaista henkilötietojen anonymisointia ei olisi hyväksytty. Ks. esim. Tietosuojavaltuutettu 25.4.2019, s. 1–2. Tietosuojavaltuutetun (TSV) mukaan siitä johtuen, että TSA:n 4(1)(1) artiklan mukainen henkilötiedon määritelmä ei jätä EU:n jäsenvaltioille kansallista liikkumavaraa, ei jäsenvaltioiden kansallisilla viranomaisilla olisi perusteita määritellä tietoja anonyymeiksi tiedoiksi, sillä tällöin kansallisella tasolla määriteltäisiin, mikä olisi henkilötietoa ja mikä ei. TSV:n tulkinnan mukaan kansallisilla lainsäätäjillä ei ole myöskään TSA:n sääntelyn perusteella toimivaltaa määritellä kansallisessa lainsäädännössä jotain tietoa anonyymiksi tiedoksi, jos se on henkilötietoa TSA:n nojalla.

[387] Merkittävää on, että DSP:n näkemyksen mukaan teknologian kehityksellä ei olisi minkäänlaista vaikutusta tietosuojalainsäädännön vaatimukset täyttävän anonymisoinnin arvioinnissa, ks. ratkaisun D.3-osan 5 kappale, jossa DPA toteaa, että ”*Der seitens des Beschwerdeführers ins Treffen geführte Umstand, dass „die Daten zu einem späteren Zeitpunkt „deanonymisiert werden könnten“, vermag daran nichts zu ändern. Eine Löschung liegt dann vor, wenn die Verarbeitung und Nutzung der personenbezogenen Daten einer betroffenen Person – so wie im vorliegenden Fall – nicht mehr möglich ist. Dass sich zu irgendeinem Zeitpunkt eine Rekonstruktion (etwa unter Verwendung neuer technischer Hilfsmittel) als möglich erweist, macht die „Löschung durch Unkenntlichmachung“ nicht unzureichend. Eine völlige Irreversibilität ist daher – unabhängig vom verwendeten Mittel zur Löschung – nicht notwendig*“. Näin ollen DPA:n tulkinnan perusteella henkilötietojen anonymisoinnin arvioinnissa tulisi ottaa huomioon vain tämänhetkinen tekniikan taso.

[388] TSA 5(1)(f).

ja näin vähentäisi rekisteröityidelle käsittelystä seuraavia tietosuojariskejä. Henkilötietojen anonymisointi ei tästä huolimatta ole aina tarkoituksenmukainen ratkaisu, sillä muun muassa erilaisessa sektorikohtaisessa lainsäädännössä edellytetään henkilötietojen säilyttämistä tunnisteellisessa muodossa tietyn ajan.[389]

Tietosuojaperiaatteiden soveltamattomuuden ohella henkilötietojen anonymisoinnilla on myös merkitystä TSA:n 6 artiklassa säädettyjen henkilötietojen käsittelyperusteiden näkökulmasta.[390] Anonymisoinnin näkökulmasta merkittävimmiksi käsittelyperusteiksi on mahdollista katsoa 6(1)(a) artiklan mukainen *suostumus* ja 6(1)(f) mukainen *oikeutettu etu*.[391] Rekisteröidyn suostumuksella tarkoitetaan TSA:n 4(1)(11) artiklan mukaan:

> *”mitä tahansa vapaaehtoista, yksilöityä, tietoista ja yksiselitteistä tahdonilmaisua, jolla rekisteröity hyväksyy henkilötietojensa käsittelyn antamalla suostumusta ilmaisevan lausuman tai toteuttamalla selkeästi suostumusta ilmaisevan toimen,”*

Suostumuksen tarkemmista edellytyksistä säädetään TSA:n 7 artiklassa,[392] jonka mukaan suostumuksen tulee olla aidosti vapaaehtoinen, sekä annettu ennalta määriteltyyn, nimenomaiseen ja lailliseen tarkoitukseen. Rekisteröity voi esimerkiksi antaa suostumuksen henkilötietojensa käsittelyyn vain tiettyä yksittäistä käyttötarkoitusta varten. Näin ollen suostumuksen luonteesta johtuen rekisterinpitäjän on pyydettävä rekisteröidyltä uutta suostumusta, jos henkilötietojen käsittelyn tarkoitus muuttuu. Tämä tekee suostumuksesta haastavan käsittelyperusteen, jos tarkoituksena on myöhemmin analysoida kerättyjä henkilötietoja alkuperäisen käyttötarkoituksen kanssa yhteensopimattomalla tavalla.[393] Suostumukseen käsittelyperusteena on toisaalta kohdistunut oikeuskirjallisuudessa huomattavaa kritiikkiä, sillä internetin kontekstissa rekisteröidyn suostumus perustuu usein pelkästään yksittäiseen klikkaukseen ilman minkäänlaisia takeita siitä, että rekisteröity olisi lukenut palvelun tietosuojaan liittyvät käyttöehdot.[394] Lisäksi varsinkin, kun rekisteröityjä on paljon tai suoraa kontaktia

[389] Esimerkiksi yritysten tulee säilyttää muun muassa kirjanpidon kannalta merkityksellisiä tietoja usein tietosuojalainsäädännön mukaisia säilytysaikoja pidempään.
[390] TSA 6(1) artiklan mukaan käsittely on lainmukaista ainoastaan, jos ja vain siltä osin kuin vähintään yksi 6(1) artiklan edellytyksistä täyttyy.
[391] Nämä ovat myös tosiasiallisesti yleisimmät rekisterinpitäjien käyttämät käsittelyperusteet.
[392] TSA 7 artikla. Tarkemmin suostumuksesta ks. WP 259 rev.01.
[393] Quinn 2017, s. 6–8.
[394] Koops 2014, s. 251–253; Kiss – Szőke 2015, s. 316–317.

rekisteröityihin ei ole, on suostumuksen pyytäminen uutta käyttötarkoitusta varten jokaiselta rekisteröidyltä erikseen joko erittäin vaikeaa, ellei mahdotonta.[395] Tästä johtuen tietyissä tilanteissa henkilötietojen anonymisointi on käytännössä ainoa vaihtoehto, jos kerättyjä tietoja halutaan hyödyntää uusiin käyttötarkoituksiin, jotka eivät ole yhteensopivia alkuperäisen käyttötarkoituksen kanssa.[396]

Toinen merkittävä käsittelyperuste on TSA:n 6(1)(f) artiklan mukainen oikeutettu etu, joka tarkoittaa henkilötietojen käsittelyä sillä perusteella, että ”*käsittely on tarpeen rekisterinpitäjän tai kolmannen osapuolen oikeutettujen etujen toteuttamiseksi, paitsi milloin henkilötietojen suojaa edellyttävät rekisteröidyn edut tai perusoikeudet ja –vapaudet syrjäyttävät tällaiset edut, erityisesti jos rekisteröity on lapsi.*”[397] Oikeutetun edun käyttäminen käsittelyperusteena edellyttää niin sanotun tasapainotestin tekemistä, jolla rekisterinpitäjän ja rekisteröidyn etuja ja oikeuksia punnitaan keskenään.[398]

3.2.3 Henkilötietojen anonymisointi henkilötietojen käsittelynä

Henkilötietojen anonymisoinnin yhteydessä herää kysymys siitä, onko henkilötietojen anonymisointi prosessina sellaista henkilötietojen käsittelyä, jonka tulisi olla yhteensopivaa sen käyttötarkoituksen kanssa, jota varten henkilötietoja on alun perin kerätty.[399] TSA:n 4(1)(2) artiklan mukaisessa henkilötietojen käsittelyn määritelmässä säädetään yksityiskohtaisesti siitä, millaiset toimet katsotaan henkilötietojen käsittelyksi. Sen mukaan henkilötietojen käsittelyllä tarkoitetaan:

> *”*[…] *toimintoa tai toimintoja, joita kohdistetaan henkilötietoihin tai henkilötietoja sisältäviin tietojoukkoihin joko automaattista tietojenkäsittelyä käyttäen tai manuaalisesti, kuten tietojen* […] *muokkaamista tai muuttamista,* […]*.”*

Henkilötietojen anonymisoinnin voidaan katsoa olevan henkilötietojen käsittelyä, sillä anonymisointiprosessissa henkilötietoja mitä ilmeisimmin muokataan tai muutetaan tai sekä että.[400] Näin ollen anonymisointi on sellaista henkilötietojen myöhempää käsittelyä, jonka tulee olla yhteensopivaa alkuperäisen käyttötarkoituksen kanssa. Tässä yhteydessä on myös tärkeää huomioida, että kyseiset

[395] Mostert et al. 2015, s. 957; Quinn 2017, s. 2.
[396] Uuden suostumuksen pyytämisen haastavuudesta jokaiselta rekisteröidyltä erikseen on kirjoitettu paljon muun muassa lääketieteellisen tutkimuksen yhteydessä. Tästä on käytetty termiä ”*consent or anonymize*” –approach. Ks. esim. Mostert et al. 2015; Sethi – Laurie 2013, s. 173–176 ja Quinn 2017, s. 2.
[397] TSA 6(1)(f) artikla.
[398] WP 217, s. 9. Tasapainotestistä tarkemmin ks. WP 217, s. 55–68, jossa on opas tasapainotestin tekemisestä ja käytännön esimerkkitilanteita.
[399] Myöhemmän käsittelyn yhteensopivuuden arvioinnista ks. tarkemmin WP 203, s. 21–23.
[400] Vastaavasti kuin henkilötietojen pseudonymisoinnissa tietoja muokataan tai muutetaan. Ks. myös IAPP 2019.

henkilötiedot tulee olla alun perin kerätty tietosuojalainsäädännön vaatimuksien mukaisesti ennen niiden anonymisoimista.[401] Koska henkilötietojen anonymisointi on henkilötietojen myöhempää käsittelyä, tulee anonymisoinnin oikeutusta lähtökohtaisesti arvioida TSA:n 6 artiklan mukaisten henkilötietojen käsittelyperusteiden kannalta.

Tietosuojatyöryhmä on tulkinnut anonymisoinnin voivan olla sellaista henkilötietojen myöhempää käsittelyä, joka on yhteensopivaa tietojen alkuperäisen käyttötarkoituksen kanssa, jos anonymisoinnin myötä henkilötiedoista tulee tosiasiallisesti anonyymejä tietoja.[402] Tietosuojatyöryhmän tulkinnalla on merkittävä vaikutus anonymisointiprosessin oikeudellisen arvioinnin kannalta, sillä tällaisen tulkinnan seurauksena henkilötietojen anonymisointi ei edellytä uutta käsittelyperustetta, kuten esimerkiksi suostumuksen hankkimista rekisteröidyiltä.[403]

Henkilötietojen käsittelyn oikeutuksen näkökulmasta toinen kysymys sen sijaan on sellainen henkilötietojen de-identifiointi[404], joka ei kuitenkaan täytä anonymisoinnin edellytyksiä. Lähtökohtaisesti tällainen tietojen de-identifiointi on sellaista henkilötietojen käsittelyä, johon rekisterinpitäjällä tulisi olla TSA:n 6 artiklan mukainen käsittelyperuste. Siitä syystä, että suostumuksen kerääminen rekisteröidyiltä erikseen voi olla käytännössä mahdotonta, on erilaisten toimijoiden intressissä käyttää de-identifioinnin käsittelyperusteena edellä mainittua *oikeutettua etua*.[405]

Käytännössä henkilötietojen de-identifiointi on hyväksyttyä tasapainotestin nojalla lähes poikkeuksetta, sillä de-identifioinnin voidaan katsoa olevan sellaista käsittelyä, joka on sekä rekisterinpitäjän että rekisteröidyn intressissä. De-identifiointi vähentää rekisterinpitäjälle henkilötietojen käsittelystä

[401] WP 216, s. 7.
[402] Ibid. Tietosuojatyöryhmä toteaa tämän anonymisointitekniikoita koskevassa lausunnossaan eksplisiittisesti: "*Accordingly, the Working Party considers that anonymisation as an instance of further processing of personal data can be considered to be compatible with the original purposes of the processing but only on condition the anonymisation process is such as to reliably produce anonymised information in the sense described in this paper*".
[403] Tästä huolimatta rekisteröidyille olisi asianmukaista vähintään ilmoittaa henkilötietojen keräämisen yhteydessä, että kyseiset tiedot tullaan todennäköisesti anonymisoimaan myöhemmin. Yleinen tietosuojalauseke tästä on esimerkiksi "your data will be used anonymously, e.g. for statistical purposes, and will not be linked back to you".
[404] De-identifiointi ei ole tietosuojalainsäädännön sellaisenaan tunnistama termi, mutta sitä käytetään muun muassa oikeuskirjallisuudessa ja erilaisissa tietosuojaoppaissa kuvaamaan kaikenlaista henkilötietojen tunnistettavuuden vaikeuttamista. De-identifioinnin pisimmälle menevänä muotona pidetään anonymisointia, mutta sen alaan kuuluu lisäksi muun muassa henkilötietojen pseudonymisointi, sekä TSA 11 artiklan mukaiset tiedot, joiden käsittely ei edellytä rekisteröityjen tunnistamista. Ks. Hintze 2018, s. 87, jonka mukaan TSA tunnistaa erilaiset de-identifioinnin muodot huomattavasti hienostuneemmalla tavalla kuin henkilötietodirektiivi.
[405] TSA 6(1)(f). Huomionarvoista on kuitenkin, että oikeutettua etua käsittelyperusteena ei kuitenkaan sovelleta tietojenkäsittelyyn, jota viranomaiset suorittavat tehtäviensä yhteydessä. Näin ollen oikeutettu etu on pätevä käsittelyperuste henkilötietojen anonymisoinnille vain yksityisen sektorin toimijoille.

aiheutuvaa riskiä, kun se samalla vahvistaa rekisteröidyn perusoikeuksien- ja vapauksien suojaa henkilötietojen käsittelyssä. Näin ollen lähtökohtaisesti useimmissa tapauksissa rekisterinpitäjillä on lainmukainen käsittelyperuste henkilötietojen de-identifioimiseksi oikeutetun edun perusteella.[406]

Eräs anonymisoinnin ja de-identifioinnin yhteydessä huomioon otettava asia on kysymys siitä, tulisiko itse anonymisointiprosessi dokumentoida. TSA:n 5(1) artiklan tietosuojaperiaatteiden yhteydessä 5(2) kohdassa säädetään osoitusvelvollisuudesta, jonka mukaan:

> *”Rekisterinpitäjä vastaa siitä, ja sen on pystyttävä osoittamaan se, että 1 kohtaa on noudatettu (”osoitusvelvollisuus”).*

Osoitusvelvollisuus tarkoittaa käytännössä sitä, että rekisterinpitäjän tulee pystyä näyttämään toteen tietosuojaperiaatteiden noudattaminen henkilötietojensa käsittelyssä.[407] Eräs tapa todistaa osoitusvelvollisuuden täyttämistä on TSA:n 35 artiklan mukaisen tietosuojan vaikutustenarvioinnin laatiminen, jossa rekisterinpitäjä kuvaa tarkasti toimintansa yhteydessä tapahtuvan henkilötietojen käsittelyn.[408] Tietosuojan vaikutustenarvioinnin laatiminen henkilötietojen anonymisoinnista voisi olla yksi keino, jolla rekisterinpitäjä osoittaa, että henkilötietojen anonymisointi on toteutettu tietosuojalainsäädännön vaatimukset täyttävällä tavalla. Tämä olisi rekisterinpitäjän kannalta erityisen hyödyllistä etenkin siinä tapauksessa, jos esimerkiksi viiden vuoden kuluttua anonymisoinnista kävisi ilmi, että tekniikan kehittymisen myötä tiedoista tulisi jälleen henkilötietoja. Käytännössä rekisterinpitäjä kuvaisi vaikutustenarvioinnissa muun muassa henkilötietojen anonymisoinnin suorittamiseksi toteuttamansa asianmukaiset tekniset ja organisatoriset toimenpiteet, anonymisointiin liittyvät riskit sekä käyttämänsä anonymisointitekniikan tai –tekniikat.[409] Vaikutustenarviointi ei ole kuitenkaan ainut keino, jolla rekisterinpitäjä toteuttaa osoitusvelvollisuuttaan. TSA:n 5(2) kohdassa käytetty sanamuoto ”on pystyttävä osoittamaan” viittaisi siihen, että rekisterinpitäjän tulee proaktiivisesti pystyä osoittamaan tietosuojaperiaatteiden noudattaminen myös niin käsittelyn lopputuloksen kuin sen toteutustavan osalta.[410]

---

[406] IAPP 2019.
[407] Korpisaari et al. 2018, s. 95–96. Osoitusvelvollisuus ei ollut oikeudellisesti rekisterinpitäjiä sitova ennen tietosuoja-asetusta. Tarkemmin osoitusvelvollisuudesta ks. WP 173.
[408] TSA 35(1) artiklan mukaan ”*Jos tietyntyyppinen käsittely etenkin uutta teknologiaa käytettäessä todennäköisesti aiheuttaa – käsittelyn luonne, laajuus, asiayhteys ja tarkoitukset huomioon ottaen – luonnollisen henkilön oikeuksien ja vapauksien kannalta korkean riskin, rekisterinpitäjän on ennen käsittelyä toteutettava arviointi suunniteltavien käsittelytoimien vaikutuksista henkilötietojen suojalle. Yhtä arviota voidaan käyttää samankaltaisiin vastaavia korkeita riskejä aiheuttaviin käsittelytoimiin”.* Tarkemmin tietosuojan vaikutustenarvioinnista ks. WP 248.
[409] Ks. kuitenkin tähän liittyvistä haasteista esim. Purtova 2018, s. 79.
[410] Ks. osoitusvelvollisuudesta myös TSA 24 artikla ja TSA:n 74 johdantokappale.

3.2.4 Anonymisointitekniikat

Tietosuojalainsäädännön näkökulmasta henkilötietojen anonymisointi on sekä oikeudellinen että tekninen ongelma. Pohjimmiltaan anonymisointi on kuitenkin ennemminkin tekninen ongelma, sillä juristit harvoin kontribuoivat oikeaoppisen anonymisoinnin tekniseen toteuttamiseen lainsäädännön asettamien vaatimusten tulkintaa ja systematisointia lukuun ottamatta. Henkilötietojen anonymisointi vaatii käytännössä sellaisia teknisiä taitoja, joita perinteinen juristikoulutus ei tarjoa, joten anonymisoinnin toteuttamiseen tarvitaan muun muassa tilastotieteilijöiden ja muiden teknisten alojen osaajien apua. Tästä huolimatta juristeja ei tule kuitenkaan unohtaa henkilötietojen anonymisointia toteuttaessa, sillä usein teknisten osaajien ja juristien näkemykset tehokkaasti toteutetusta anonymisoinnista saattavat erota toisistaan merkittävästi.[411] Teknisen alan henkilön pitäessä tiettyä tietomassaa anonymisoituna, näkee juristi samat tiedot pikemminkin pseudonymisoituina, ja kuten edeltä käy ilmi, tietosuojalainsäädäntö suhtautuu pseudonymisoituihin tietoihin ja anonyymeihin tietoihin hyvin eri tavoin.[412]

Anonymisoinnin teknisestä toteuttamisesta ei ole tällä hetkellä EU:ssa ajantasaisia ohjeita osittain siitä syystä, että anonymisointitekniikat kehittyvät hyvin nopeasti.[413] Anonymisointitekniikoiden tehokkuutta arvioitaessa on huomattava, että myös anonymisoinnin purkamiseen soveltuvat teknologiat kehittyvät jatkuvasti, ja rekisterinpitäjien on huomioitava tämä omaan kontekstiin parhaiten soveltuvaa anonymisointiteknologiaa valitessaan.[414] Tällä hetkellä eurooppalaisen tietosuojalainsäädännön mukaisen anonymisoinnin tulkinnan näkökulmasta ensisijaisena lähteenä toimii tietosuojatyöryhmän

[411] Muun muassa yhdysvaltalaiset tietotekniikan tutkijat Aloni Cohen ja Kobbi Nissim argumentoivat huhtikuussa 2019 julkaistussa artikkelissaan Towards Formalizing the GDPR's Notion of Singling Out, että: "*There is a significant conceptual gap between legal and mathematical thinking around data privacy. The effect is uncertainty as to which technical offerings adequately match expectations expressed in legal standards*". Kyseinen artikkeli on siitä merkittävä, että sen toisena kirjoittajana on Georgetownin yliopiston tietotekniikan professori Kobbi Nissim, jota pidetään yhtenä differentiaalisen yksityisyyden kehittäjistä. Differentiaalisesta yksityisyydestä tarkemmin ks. esim. WP 216, s. 15–16. Ks. myös juristien ja teknisten osaajien poikkeavista näkemyksistä anonymisointiin liittyen Purtova 2018, s. 73.
[412] TSA 26 johdantokappale. Ks. toisaalta Mourby et al. 2018, s. 223–224.
[413] Ks. DPC Guidance 2019, jossa Irlannin tietosuojaviranomainen toteaa, että parhaillaan on käynnissä paljon anonymisointiin liittyvää tutkimusta ja tietämys erilaisten anonymisointitekniikoiden tehokkuudesta muuttuu koko ajan. Tästä syystä on käytännössä mahdotonta todeta varmasti, että jokin anonymisointitekniikka suojelee rekisteröityjä tunnistamiselta sadan prosentin varmuudella. Ks. myös ajantasaisempina ohjeina anonymisointitekniikoista PDPC 2018, joissa Singaporen tietosuojaviranomaisen tarkastelee ja arvioi erilaisia anonymisointitekniikoita. Kyseisistä ohjeista joista voi olla hyötyä arvioitaessa valitun anonymisointitekniikan tehokkuutta, mutta ohjeet ovat kuitenkin avuksi vain anonymisoinnin teknisen puolen tarkastelussa, sillä Singaporen kansallinen tietosuojalainsäädäntö eroaa TSA:n sääntelystä.
[414] Anonymisoinnin takaisinmallintamisesta ("*reverse-engineering*") on useita esimerkkejä, ks. esim. Guardian 13.7.2018.

anonymisointitekniikoita koskeva lausunto 05/2014 (WP 216). Kyseisessä lausunnossa tietosuojatyöryhmä on muun muassa todennut, että tunnistamisen mahdollistavan teknologian nopean kehittymisen seurauksena todellisen anonymisoinnin saavuttaminen on erittäin vaikea tehtävä.[415]

Tietosuojatyöryhmä on lausunnossaan arvioinut erilaisia anonymisointitekniikoita kolmen kriteerin valossa, jotka on esitetty kysymysten muodossa. Keskeisimmät anonymisointitekniikoiden tehokkuutta arvioitaessa huomioon otettavat kriteerit ovat:

> *”(i) onko yksilö edelleen mahdollista erottaa joukosta,*
>
> *(ii) onko tietojen yhdistäminen yksilöön edelleen mahdollista, ja*
>
> *(iii) voidaanko yksilöä koskevat tiedot päätellä. ”.*[416]

Tietosuojatyöryhmä korostaa, että kullakin tekniikalla on erilaisia vahvuuksia ja heikkouksia, ja kussakin tilanteessa tulisi erikseen arvioida mikä anonymisointimenetelmä soveltuisi tilanteeseen parhaiten.[417] Lausunnossa myös tiedostetaan, että oikein toteutettu tietosuojalainsäädännön vaatimukset täyttävä anonymisointi voi olla sekä julkisille että yksityisille toimijoille toimiva strategia, jonka avulla tiedosta saatavat hyödyt säilyvät ja tietosuojariskit vähenevät.[418] Tästä huolimatta on syytä erikseen korostaa, että henkilötietojen tehokas anonymisointi on kaikkea muuta kuin yksinkertainen tehtävä. Jos rekisterinpitäjän tarkoitus on anonymisoida tiedot siten, ettei anonymisoiduissa tiedoissa ole tilastollisesti merkitsevää eroa[419] alkuperäisiin tietoihin verrattuna, rekisteröityjen uudelleentunnistamisen riski on huomattavasti suurempi verrattuna tilanteeseen, jossa tiedoista pyritään vain poistamaan tunnistettavuus.[420] Tavanomaisesti henkilötiedot pyritään anonymisoimaan tällä tavalla, sillä tällöin niiden hyödyntäminen uusin käyttötarkoituksiin data-analytiikan menetelmin on vielä mahdollista.[421]

---

[415] WP 216, s. 3.
[416] Ibid.
[417] WP 216, s. 23.
[418] WP 216, s. 3.
[419] Tilastollisesti merkitsevä ero (”*Statistical significant difference*”) on tilastotieteessä käytetty termi, joka kuvaa tietoaineiston laatua. Lähtökohtaisesti anonymisoiduissa tiedoissa ei tulisi olla tilastollisesti merkitsevää eroa, jotta tietoja voisi hyödyntää erilaisissa tutkimuksissa vielä anonymisoinnin jälkeenkin.
[420] Ks. esim. Ohm 2009, s. 1703–1704.
[421] ICO 2017b, s. 58. Ks. Myös Hintze 2018, s. 87 ja Stalla-Bourdillon – Knight 2016, s. 285.

Jos rekisterinpitäjän tarkoituksena on ainoastaan poistaa tunnistettavuus tiedoista, eikä käyttää tietoja enää uusiin käyttötarkoituksiin, on anonymisointi mahdollista suorittaa suhteellisen suoraviivaisesti poistamalla tiedoista riittävästi henkilöiden tunnistettavuuden mahdollistavia elementtejä.[422] Tällaista tunnistettavuuden poistamista kutsutaan karkeistamiseksi, eli *aggregoinniksi.*[423] Tietosuojatyöryhmän mukaan esimerkiksi organisaatio, joka kerää tietoja yksittäisten ihmisten yksittäisistä matkoista, voi anonymisoida kyseiset tiedot tietosuojalainsäädännön vaatimukset täyttävällä tavalla aggregoimalla tiedot.[424] Tällainen karkeistamalla toteutettu anonymisointi olisi mahdollista toteuttaa jättämällä tiedoista jäljelle esimerkiksi ainoastaan tiedon siitä, että perjantai-iltaisin tietyllä reitillä on tavanomaisesti kaksinkertainen määrä matkustajia torstai-iltaan verrattuna.[425] Aggregoinnin yhteydessä käsittelemättömät tiedot tulisi poistaa, että tiedot olisivat rekisterinpitäjän näkökulmasta anonyymejä tietoja. Käytännössä aggregoimalla tehty anonymisointi mahdollistaa sen, että rekisterinpitäjä voi tietosuojalainsäädännön estämättä esimerkiksi myydä tai luovuttaa kyseisiä anonymisoituja tietoja kolmansille osapuolille, sillä tällaisista tiedoista ei ole mahdollista erottaa yksilöitä.[426]

Tietosuojatyöryhmä on anonymisointitekniikoita koskevassa lausunnossaan jakanut anonymisointiteknologiat kahteen tekniikkaryhmään, jotka ovat 1) *satunnaistaminen* ja 2) *luokituksen karkeistaminen.*[427] Lausunnossa tekniikoiden toimivuutta arvioidaan edellä mainittujen kolmen kriteerin valossa, jotka ovat henkilön tunnistaminen *erottamalla*, *yhdistämällä* tai *päättelemällä.*[428] Nämä tietosuojatyöryhmän muodostamat kriteerit muodostavat samalla riskit, joiden valossa henkilötietojen anonymisointia tulisi eurooppalaisessa tietosuojalainsäädännössä arvioida.

[422] WP 216, s. 5.
[423] WP 216, s. 9. Myös aggregointi on tilastotieteessä käytetty termi, jolla tarkoitetaan eri attribuuttien, eli tiedoissa olevien arvojen karkeistamista siten, ettei tiedoista ole enää mahdollista eristää yksittäisiä arvoja, joista yksilöitä voisi olla mahdollista tunnistaa.
[424] WP 216, s. 9.
[425] Ibid.
[426] Organisaatioiden tulisi kuitenkin huomioida myös anonymisoitujen tietojen hyödyntämisen psykologinen vaikutus ihmisiin, sillä anonymisoitujen tietojen käyttäminen saattaa aiheuttaa ihmisille hyväksikäytetyn olon. Tunnettu esimerkki tästä on GPS-satelliittien signaaliin ja sähköisiin karttoihin perustuvia paikannuslaitteita kehittävän hollantilaisen yrityksen TomTomin aiheuttama kohu. TomTom oli myynyt asiakkailtaan keräämiä ja aggregoimalla anonymisoituja ajo- ja liikennetietoja hollantilaisille viranomaisille, joita viranomaiset olivat käyttäneet nopeusrajoitusvalvonnan tehostamiseen sijaintitietojen perusteella. Tapauksesta seurasi suuri kohu, sillä ihmiset muun muassa epäilivät anonymisoinnin tehokkuutta. Lopulta Hollannin tietosuojaviranomainen (Autoriteit Persoonsgegevens, AP) tutki tapauksen ja päätyi toteamaan, että TomTom oli toiminut silloisen tietosuojalainsäädännön mukaisesti myydessään tietoja viranomaisille. Yhtiö ei ollut kuitenkaan pyytänyt riittävän täsmällistä suostumusta asiakkailtaan sijaintitietojen käsittelemiseen, jonka seurauksena TomTom päivitti tietosuojaselostettaan. Ks. tarkemmin tapauksesta Reuters 12.1.2012.
[427] WP 216, s. 12.
[428] Ibid, s. 11–12.

Satunnaistaminen on joukko erilaisia tekniikoita, joiden avulla tietojen totuudenmukaisuutta pyritään heikentämään siten, ettei tietoja voisi olla enää mahdollista yhdistää tiettyyn henkilöön. Näiden tekniikoiden hyödyntämisen seurauksena voidaan teoriassa poistaa tietojen yhteys tiettyyn yksilöön. Satunnaistamistekniikoilla pyritään käytännössä siihen, että tietojen tilastollinen ainutlaatuisuus säilyy, mutta yksittäisten henkilöiden tunnistaminen tiedoista ei ole enää mahdollista. Tämä tarkoittaa sitä, että anonymisoiduissa tiedoissa on lähtökohtaisesti edelleen samat tilastolliset ominaisuudet, mutta tiedoissa oleviin yksittäisiin henkilöihin liittyvät tiedot eivät ole enää totuudenmukaisia.[429] Tietosuojatyöryhmän mukaan satunnaistamisteknologiat suojaavat rekisteröityjä tunnistamiselta, mutta ne eivät useimmissa tapauksissa ole yksinään riittäviä tietosuojalainsäädännön vaatimukset täyttävän anonymisoinnin toteuttamiseksi.[430]

Yleinen satunnaistamistekniikka on *kohinan lisääminen*, joka tarkoittaa käytännössä tarkkuuden vähentämistä tiedoista. Tiedoissa on tavanomaisesti erilaisia attribuutteja, jotka tarkoittavat tiedoissa olevia arvoja, kuten henkilöiden pituuden laskemista senttimetrin tarkkuudella. Kohinan lisääminen tarkoittaa tiedoissa olevien arvojen muuttamista siten, ettei yksittäisiä henkilöitä olisi enää mahdollista tunnistaa tiedoista. Esimerkiksi henkilöiden pituuden ilmaiseminen +-5 senttimetrin tarkkuudella yhden senttimetrin tarkkuuden sijaan vaikeuttaa yksittäisten henkilöiden tunnistamista tiedoista. Tällainen kohinan lisääminen tarkoittaisi käytännössä sitä, että jokainen 180 senttimetriä pitkä henkilö ilmoitettaisiin tiedoissa 175-185 senttimetrin pituisena. Tällainen kohinan lisääminen vähentäisi tietojen tarkkuutta tiedoissa olevien henkilöiden osalta siten, että heidän tunnistamisensa pituuden perusteella vaikeutuisi.[431]

Toinen yleinen satunnaistamistekniikka on *permutaatio*, jolla tarkoitetaan attribuuttien siirtämistä siten, että arvojen ja yksilöiden välinen korrelaatio häviäisi. Permutaatiolla tarkoitetaan tekniikkaa, jonka avulla tiedoissa olevia arvoja pyritään vaihtamaan keskenään siten, että yksittäisten henkilöiden tunnistaminen ei olisi enää mahdollista. Tämän yhteydessä on kuitenkin syytä kiinnittää huomiota siihen, että keskenään vaihdettavat attribuutit liittyvät toisiinsa, sillä attribuuttien välisen loogisen suhteen katkeaminen voisi mahdollistaa permutaatiossa käytetyn logiikan tunnistamisen siten, että prosessi olisi mahdollista peruuttaa.[432]

[429] WP 216, s. 12.
[430] Ibid.
[431] Tässä tulisi kuitenkin huomata tilaston otanta: jos tiedoissa on vain yksi 150 senttimetriä pitkä henkilö, hän voisi olla tunnistettavissa kohinan lisäämisestä huolimatta.
[432] Permutaatiosta tekniikkana tarkemmin, ks. esim. Journal of Intelligent Information Systems –lehdessä joulukuussa 2016 julkaistu artikkeli, Li et al. 2016, jossa käsitellään yksityiskohtaisesti permutaatiota teknisestä näkökulmasta.

Kolmas ja samalla kenties tunnetuin satunnaistamistekniikka on *differentiaalinen yksityisyys*, joka poikkeaa merkittävästi lähestymistavaltaan sekä kohinan lisäämisestä että permutaatiosta.[433] Differentiaalisella yksityisyydellä tarkoitetaan erilaisten satunnaistamiseen perustuvien anonymisointitekniikoiden sekä teknisten ja organisatoristen toimenpiteiden yhdistelmää, joiden avulla rekisterinpitäjä muodostaa hallussaan olevista tiedoista anonymisoituja näkymiä. Tällaisen tekniikan hyödyntämisen tarkoituksena on, että esimerkiksi julkinen viranomainen voisi luovuttaa hallussaan olevista tietovarannoista anonymisoituja palasia erilaisille toimijoille muun muassa tieteellistä tutkimusta varten.[434] Siitä syystä, että rekisterinpitäjä säilyttää alkuperäiset tiedot, joista se toimintansa yhteydessä luovuttaa anonymisoituja osia, ovat tiedot edelleen vähintäänkin rekisterinpitäjän näkökulmasta henkilötietoja.

Differentiaalisen yksityisyyden tehokas ja oikeaoppinen toteuttaminen edellyttää sitä, että rekisterinpitäjän tulee pitää tarkkaa kirjaa siitä, millaisia tietoja erilaisille toimijoille on luovutettu. Tämä johtuu siitä, että rekisterinpitäjän aikaisemmin luovuttamat tiedot voi olla mahdollista yhdistää rekisterinpitäjän myöhemmin lähettämiin tietoihin yksittäisten henkilöiden tunnistamisen mahdollistavalla tavalla.[435] Riski tietojen yhdistämisestä aiheuttaa sen, että differentiaalista yksityisyyttä ei tulisi käyttää anonymisointitekniikkana avointen hakukoneiden yhteydessä, sillä niiden yhteydessä hakukoneen käyttäjien tekemiä tietopyyntöjä ei ole mahdollista valvoa riittävällä tavalla.[436]

Toinen yleinen anonymisointitekniikkaryhmä on tietojen *luokituksen karkeistaminen*, jolla tarkoitetaan tunnistettavuuden poistamista tiedoista yleistämällä tietoja.[437] Käytännössä tämä tarkoittaa tiedoissa olevien attribuuttien mittakaavan tai suuruusluokan muuttamista. Tavanomaisesti tietojen mittakaavaa tai suuruusluokkaa on mahdollista muuttaa esimerkiksi käsittelemällä alueita kaupunkien sijaan tai vuosia kuukausien sijaan. Muun muassa edellä yksinkertaisesti kuvattu tietojen aggregointi kuuluu luokituksen karkeistamiseen perustuvien anonymisointitekniikoiden ryhmään, mutta on ole-

[433] WP 216, s. 15.
[434] Tähän liittyen, ks. HE 159/17 vp., jossa säädetään Suomessa 1.5.2019 voimaantulleesta sosiaali- ja terveystietojen toissijaisesta käytöstä annetusta laista (552/2019), jota kutsutaan yleisesti toisiolaiksi. Kyseisessä laissa säädetään vaiheittain perustettavasta Tietolupaviranomaisesta, jolla on valtuudet luovuttaa aggregoitua tilastotietoa erilaisille toimijoille muun muassa kehitys ja innovaatiotoimintaa varten.
[435] Kounadi et al. 2018, s. 7.
[436] WP 216, s. 15.
[437] Ibid, s. 16.

massa myös hienostuneempia luokituksen karkeistamiseen perustuvia teknologioita. Tietosuojatyöryhmä esittelee anonymisointitekniikoita koskevassa lausunnossaan karkeistamistekniikkoina k-anonymiteetin ja l-diversiteetin, jotka ovat teknisiä ja hienostuneita anonymisointitekniikoita.[438]

Edellä käsiteltyihin anonymisointitekniikoihin sisältyy erilaisia vahvuuksia ja heikkouksia, jonka lisäksi niitä ei ole mahdollista tarkastella yhtenä homogeenisenä tekniikkojen joukkona.[439] Osittain tästä syystä kuhunkin kontekstiin parhaiten soveltuvimman anonymisointitekniikan määrittäminen perustuu samalla tapaa kontekstisidonnaisiin seikkoihin kuin tunnistettavuuden arvioiminen ylipäänsä. Rekisterinpitäjän olisi näin ollen perusteltua anonymisointitekniikan valintaan liittyvää päätöstä tehdessään arvioida anonymisoitavia henkilötietoja kaikkien edellä 2.2 kappaleessa seikkaperäisesti käsiteltyjen henkilötiedon käsitteen osatekijöiden näkökulmasta, sekä esimerkiksi laatia prosessista TSA:n 35 artiklan mukainen tietosuojan vaikutustenarviointi.

Tietosuojalainsäädännön näkökulmasta yksi tärkeimmistä huomioista anonymisointitekniikoihin liittyen on, että nykykäsityksen mukaan käytännössä ainut tapa poistaa tunnistettavuus tiedoista varmuudella ja riskeittä on perusteellisesti toteutettu tietojen aggregoiminen, jonka seurauksena tiedoista on jäljellä enää vain aggregoitua tilastotietoa.[440] Aggregoitujenkin tietojen yhteydessä on myös syytä ottaa huomioon tietoihin mahdollisesti sisältyvät erittäin poikkeukselliset attribuutit, jotka tulee poistaa tiedoista aggregoinnin yhteydessä yksittäisinä ilmiselvinä tunnisteina.[441] Aggregoinnilla toteutettuun anonymisointiin liittyy kuitenkin se merkittävä ongelma, että tällöin tietojen hyödynnettävyys uusiin käyttötarkoituksiin heikkenee perustavanlaatuisesti, ja muun muassa tietojen tutkimuskäyttö ei ole enää lähtökohtaisesti mahdollista. Tästä syystä anonymisointiteknologia on nykyisin tietosuojaan liittyvä merkittävä liiketoiminnan ala, kun erilaiset toimijat pyrkivät kehittämään sellaisia anonymisointitekniikoita, joiden avulla tietojen hyödynnettävyys säilyy, mutta luonnollisten henkilöiden tunnistettavuus poistuu.[442]

---

[438] En kuitenkaan käsittele tutkielmassa yksityiskohtaisemmin monimutkaisia tilastotieteellisiä anonymisointimenetelmiä, sillä niiden tarkastelu ei ole tarkoituksenmukaista lainopillisen tutkielman yhteydessä. Tästä huolimatta on syytä todeta, että k-anonymiteetti on lukuisissa yhteyksissä todettu puutteelliseksi anonymisointitekniikaksi laajojen tietoaineistojen anonymisoinnissa. Ks tähän liittyen Narayan – Shmatikov 2010, s. 2–3 ja 14.

[439] WP 216, s. 11.

[440] Ks. esim. Mayer-Schönberger – Cukier 2013, s. 154–156, joiden mukaan tehokas anonymisointi ei ole enää mahdollista big datan aikakaudella.

[441] Jos esimerkiksi rekisterinpitäjä on muodostanut tilaston tietyllä bussireitillä tiettyinä kellonaikoina matkustaneista henkilöistä ja vain yksi henkilö on matkustanut aamun ensimmäisellä vuorossa, tulee tämä tieto poistaa, sillä kyseinen henkilö on tosiasiallisesti tunnistettavissa tällaisesta tilastosta.

[442] Anonymisointitekniikoiden alalle on myös ominaista kilpajuoksu anonymisoinnin purkamiseen erikoistuneiden teknologioiden kanssa.

Tämä tarkoittaa samalla sitä, että henkilötietojen anonymisointi ei ole kertaluontoinen prosessi, vaan anonymisoinnin tehokkuutta, toisin sanoen tietojen tunnistettavuutta, tulee arvioida säännöllisesti. Näin ollen henkilötietojen anonymisointi tulisi nähdä pikemminkin vastaavana prosessina kuin tietoturva-auditoinnit, joita suoritetaan esimerkiksi säännöllisesti vuosittain. Rekisterinpitäjä voisi arvioida anonymisointia muun muassa uusien julkisesti julkaistujen tietojen sekä tunnistamisen mahdollistavien teknologioiden kehityksen näkökulmasta. Teknologisen kehityksen huomioiminen tietosuojalainsäädännön mukaisen anonymisoinnin arvioinnissa johtaa siihen, että rekisterinpitäjän tulisi olla äärimmäisen varovainen anonymisoitujen tietojen avoimessa julkaisemisessa, sillä tällaisissa tapauksissa tietojen tunnistettavuus ei ole enää rekisterinpitäjän hallittavissa.[443] Käytännössä ainut varma tapa julkaista anonymisoituja tietoja avoimesti on tietojen julkaiseminen aggregoituna tilastotietona.

3.2.5 Henkilötietojen anonymisoinnin haasteet

Henkilötietojen anonymisointiin liittyy useita haasteita, jotka johtuvat useimmiten siitä, että varsinkin suurten tietomassojen anonymisointi on todellisuudessa erittäin vaikeaa, jos tietojen hyödynnettävyys uusiin käyttötarkoituksiin halutaan säilyttää.[444] Kynnys tehokkaasti ja riittävästi toteutetulle henkilötietojen anonymisoinnille on asetettu eurooppalaisessa tietosuojalainsäädännössä erittäin korkealle, sillä henkilötietojen suoja on nykyisin tärkeä perusoikeus Euroopan unionissa.[445] Edellä tarkastellut anonymisointitekniikat ovat tehokkaita keinoja henkilötietojen tietoturvan parantamiseksi ja niiden de-identifioimiseksi, mutta usein sellaisenaan riittämättömiä tunnistettavuuden poistamiseksi tiedoista tietosuojalainsäädännön edellyttämällä tavalla. Tämä johtuu siitä, että käytännössä on aina olemassa riski yksittäisen henkilön tunnistamisesta myös anonymisoiduista tiedoista joko erottamalla hänet joukosta, yhdistämällä tiedot toisiin tietoihin tai päättelemällä.[446] Yksi suurimmista syistä tähän on, että yhä useammin yksittäinen henkilöön liittyvä *tunniste* on yhdistettävissä tunnistettavissa olevaan henkilöön, kun kyseistä tunnistetta analysoi saatavilla olevia julkisia ja ei-julkisia tietoja vasten.[447]

[443] Ks. myös julkisen sektorin avoimen datan julkaisemiseen liittyen Zuiderveen Borgesius et al. 2015, s. 2088–2091, jossa kirjoittavat arvioivat avoimen datan julkaisemiseen liittyviksi riskeiksi pelotevaikutuksen, ihmisten omiin henkilötietoihinsa kohdistuvan kontrollin katoamisen ja riskin, että avointa dataa voitaisiin käyttää henkilöiden eriarvoiseen kohteluun.
[444] Kenties kuuluisimman henkilötietojen anonymisointiin kohdistuneen vastalauseen ”*data can be either useful or perfectly anonymous but never both.*” on todennut Paul Ohm 2009, s. 1704.
[445] Lindroos-Hovinheimo 2018, s. 52.
[446] WP 216, s. 23–24; Jändel 2014, s. 48.
[447] Mayer-Schönberger – Cukier 2013, s. 154.

Tietosuojalainsäädännön mukaisen anonymisoinnin saavuttaminen edellyttää sen arvioimista, mitä kaikkea tiedoissa olevien henkilöiden tunnistamisen mahdollistavaa muuta tietoa on, tai voisi ylipäänsä olla olemassa.[448] Tämä tekee henkilötietojen anonymisoinnista haasteellista, sillä rekisterinpitäjän on käytännössä erittäin vaikea arvioida olemassa olevia julkisia tietoja tästä näkökulmasta, ei-julkisista tiedoista puhumattakaan. Lisäksi on käytännössä mahdotonta selvittää, mitkä kaikki toimijat pääsevät ylipäänsä käsiksi henkilöiden tunnistamisen potentiaalisesti mahdollistavaan tietoon ja kuinka todennäköisesti tällaisia muita tietoja tultaisiin käyttämään henkilöiden uudelleentunnistamiseksi anonymisoiduista tiedoista.[449]

Lukuisat käytännön esimerkit osoittavat, että aidosti anonyymin tietoaineiston luominen henkilötiedoista on rekisterinpitäjälle haastava tehtävä, jos aikomuksena on säilyttää tietojen hyödynnettävyys myöhempiä käyttötarkoituksia varten.[450] Useissa julkisuuteen tulleissa tapauksissa avoimesti julkaistut anonymisoidut tiedot on ollut lopulta mahdollistaa muihin olemassa oleviin ja useimmiten julkisiin tietoihin sellaisella tavalla, että tiedoista on ollut mahdollista tunnistaa yksittäisiä henkilöitä. Näissä tapauksissa ei ole ollut tietosuojalainsäädännön näkökulmasta itse asiassa kyse anonymisoiduista tiedoista, sillä tiedot ovat olleet pikemminkin vain pseudonymisoituja tietoja. Näin ollen tiedot ovat olleet koko ajan henkilötietoja ja rekisterinpitäjä on mahdollisesti julkaissut henkilötiedot ilman asianmukaista käsittelyperustetta.

Kuuluisin esimerkki tällaisesta epäonnistuneesta ja huonosti toteutetusta anonymisoinnista on Netflixin vuonna 2006 järjestämän Netflix Price kilpailun yhteydessä julkaisemat 100 miljoonaa elokuva-arvostelua 500 tuhannelta käyttäjältä.[451] Netflix oli pyrkinyt anonymisoimaan tiedot siten, että julkaistuista arvosteluista oli poistettu henkilökohtaiset tiedot, jonka lisäksi käyttäjien nimet oli korvattu satunnaisilla numeroilla. Arvostelujen julkaisemisen tarkoituksena oli, että tutkijat ympäri maailman tarkastelisivat elokuva-arvosteluja yhdessä Netflixin suositusalgoritmin kanssa ja kehittäisivät suositusalgoritmia miljoonan dollarin palkkion toivossa. Kuitenkin jo seuraavan vuoden aikana Texasin yliopiston tutkijat Arvid Narayan ja Vitaly Shmatikov onnistuivat de-anonymisoimaan osan Netflixin julkaisemista tiedoista ja tunnistamaan arvosteluja antaneita käyttäjiä.[452] Tutkijat onnistuivat de-anonymisoinnissa, kun he yhdistivät käyttäjien antamat elokuva-arvostelut ja niiden aikaleimat IMDb:stä

[448] Esayas 2015, s. 3.
[449] ICO 2012, s. 18. Tämän arvioinnissa ICO:n kehittämä motivated intruder -testi on erittäin hyödyllinen.
[450] ENISA 2015, s. 27.
[451] Mayer-Schönberger – Cukier 2013, s. 155.
[452] Narayan – Shmatikov 2008.

(*The Internet Movie Database*) löydettäviin käyttäjien antamiin julkisiin elokuva-arvosteluihin, jotka vastasivat Netflixin julkaisemissa tiedoissa olevia arvosteluja.

Toinen julkisuuteen noussut esimerkki puutteellisesti toteutetusta anonymisoinnista oli yhdysvaltalaisen internet-palveluntarjoaja AOL:n vuonna myös vuonna 2006 julkaisemat käyttäjiensä hakusanalokit.[453] Julkaistuihin tietoihin sisältyi 20 miljoonaa verkkohakua yli 650 tuhannelta käyttäjältä, jotka olivat käyttäneet AOL:n tarjoamaa verkkoportaalia. Jokainen AOL:n julkaisema verkkohaku oli yhdistettävissä jokaiselle palvelun käyttäjälle annettuun ainutlaatuiseen numerotunnisteeseen. Jo samana vuonna The New York Times -sanomalehden toimittajat onnistuivat tunnistamaan julkaistujen hakusanalokien perusteella käyttäjän nro. 4417749 Georgian osavaltiossa asuvaksi 62-vuotiaaksi leskeksi Thelma Arnoldiksi. AOL:n esimerkki osoittaa, ettei pelkän nimen korvaaminen pseudonyymillä ole riittävä toimenpide anonymisoinnin toteuttamiseksi, jos julkaistuissa tiedoissa on muita yksilön tunnistamisen mahdollistavia tietoja.[454]

Kolmas kuuluisa esimerkki puutteellisesti toteutetusta anonymisoinnista liittyy sijaintitietojen anonymisointiin. Tapauksessa oli kyse New Yorkin kaupungin vuonna 2014 avoimesti julkaisemista 173 miljoonan taksimatkan tiedoista. Julkaistuihin tietoihin oli jätetty jokaisen yksittäisen taksimatkan alkamis- ja päättymisajat ja sijainnit sekä väitetysti anonymisoituna taksin rekisterinumero ja kuljettajan identifioiva numero. Pian tietojen julkaisemisen jälkeen kävi kuitenkin ilmi, että taksikuljettajien tiedot oli salattu MD5-algoritmillä, joka oli mahdollista peruuttaa suhteellisen vähällä vaivalla, kun tiesi miltä tiedot näyttivät ennen niiden salaamista.[455]

Taksimatkoista kerättyjen tietojen julkaisemiseen liittyi kuitenkin vielä huomattavasti suurempi tietosuojariski, sillä yksittäisten taksimatkojen tietoja analysoimalla oli mahdollista tunnistaa myös huomattava määrä yksittäisiä matkustajia. Esimerkiksi vertaamalla paparazzien ottamia kuvia julkisuuden henkilöistä astumassa taksiin ja julkaistuja tietoja taksimatkoista oli muun muassa mahdollista selvittää, että mihin kukin julkisuuden henkilö matkusti.[456] Käytännössä New Yorkin kaupungin jul-

[453] New York Times 9.8.2006.
[454] Hintze 2018, s. 90. Tällä esimerkillä on kuitenkin toinen puoli: ilmeisesti vuoteen 2019 mennessä vain 5 henkilöä oli tunnistettu AOL:n julkaisemista 650 tuhannen käyttäjän tiedoista. Tämä on 0.001 prosenttia käyttäjistä, joten ehdoton enemmistö käyttäjistä pysyi anonyymeinä.
[455] Käytännössä sen perusteella, että New Yorkin taksinumerot olivat joko kuusinumeroisia tai seitsemännumeroisia, jolloin ne alkoivat numerolla 5, oli mahdollista rajata mahdollisten taksinumeroiden määrä kolmeen miljoonaan. Vastaavalla logiikalla kuljettajien numeroita oli mahdollista laskea olevan vain 22 miljoonaa. Vuonna 2014 kuljettajien tietojen deanonymisoimiseen raportoitiin kuluvan yhteensä vain tunti.
[456] Guardian 27.6.2014.

kaisemia taksimatkoja koskevia tietoja analysoimalla oli mahdollista selvittää huomattava määrä yksittäisiä henkilöitä koskevia tietoja, kun tiedot yhdisti pelkästään julkisissa lähteissä oleviin muihin tietoihin.[457]

Vuoden 2019 elokuuhun mennessä julkisuuteen on noussut yksi EU:n jäsenvaltion tietosuojaviranomaisen TSA:n nojalla antama ratkaisu koskien puutteellisesti toteutettua henkilötietojen anonymisointia. Tanskan tietosuojaviranomainen (*Datatilsynet, DPA*) antoi maaliskuussa 2019 ratkaisun, jossa se totesi taksiyhtiö Taxa 4x35:n menetelleen tietosuojalainsäädännön vastaiseksi, sillä yhtiö oli säilyttänyt asiakkaidensa henkilötietoja tarkoituksenmukaista kauemmin ja näin toiminut tietojen minimointi -periaatteen vastaisesti.[458] Yhtiö oli poistanut asiakkaiden nimet ja kotiosoitteet kahden vuoden säilytysajan jälkeen, mutta säilyttänyt asiakkaiden puhelinnumeroja vielä ylimääräiset kolme lisävuotta yhdessä taksimatkoista kerätyn tiedon kanssa.[459] Taksimatkoihin liittyen yhtiö keräsi tietoa muun muassa yksittäisten matkojen alkamis- ja päättymisosotteista. DPA:n ratkaisun perusteella yhtiön suorittama henkilötietojen anonymisointi niiden poistamisen toteuttamiseksi oli puutteellisesti toteutettu, sillä yhtiön järjestelmässä olevat rekisteröityihin liittyvät tiedot olivat edelleen yhdistettävissä tunnistettavissa oleviin henkilöihin säilytettyjen puhelinnumeroiden perusteella.[460]

[457] Fast Company 10.2.2014.
[458] Datatilsynet 2019.
[459] Menettelynsä oikeutukseksi yhtiö vetosi muun muassa siihen, että yhtiön senhetkinen IT-järjestelmä ei mahdollistanut tietojen anonymisointia riittävällä tasolla. Tässä yhteydessä DPA erikseen totesi, että tietosuojasääntöjen noudattamatta jättämistä ei voi perustella IT-järjestelmien puutteilla.
[460] DPA ehdotti yhtiölle sen menettelystä 1.2 miljoonan kruunun sakkoa, joka vastaa noin 160 tuhatta euroa ja on arviolta 2,8 prosenttia Taxa 4x35:n kansainvälisestä liikevaihdosta. Tanskalaisten tuomioistuinten tulee vielä hyväksyä DPA:n sakko, mutta tästä huolimatta DPA:n sakkoehdotus on oiva esimerkki hallinnollisesta sakosta, joka yksittäiselle yritykselle saattaa langeta maksettavaksi puutteellisesti toteutetun anonymisoinnin johdosta. Tietosuojalainsäädännön vastaisesta menettelystä seuraavista sanktioista ks. TSA 83 artikla.

4 Johtopäätökset

Tutkielmassa on tarkasteltu sekä henkilötiedon käsitteen sisältöä eurooppalaisessa tietosuojalainsäädännössä että kolikon kääntöpuolena olevaa anonyymien tietojen konseptia. Tämä henkilötietojen suojaa koskevassa lainsäädännössä omaksuttu kahtiajako on erityisen tärkeä, sillä tietojen asemoiminen henkilötiedoiksi johtaa TSA:n soveltumiseen kyseisten tietojen käsittelyssä. Lisäksi tutkielmassa on analysoitu henkilötietojen anonymisointia, joka on ennen kaikkea mahdollista nähdä prosessina, jonka seurauksena käsiteltävät tiedot siirtyvät tietosuojalainsäädännön soveltamisalan ulkopuolelle tunnistettavuuden poistamisen seurauksena.

Eurooppalaisen tietosuojalainsäädännön mukainen henkilötiedon käsite on erittäin laaja, kun sen määritelmän mukaisesti henkilötiedoiksi tulkitaan kaikki tiedot, jotka liittyvät tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön. Määritelmän laaja tulkinta on perusteltua siitä näkökulmasta, että tietosuojalainsäädäntö on pohjimmiltaan perusoikeusjuridiikkaa, jonka tulkinnassa muut tietojen käyttöön liittyvät intressit ovat toissijaisia yksilön perusoikeuksien toteutumisen kannalta.[461] Henkilötiedon käsitteen laajuus ei ole kuitenkaan aivan ongelmatonta, sillä henkilötiedon avoimeksi kirjoitettu määritelmä ja käsitteen alati laajentuva tulkinta ovat omiaan vaarantamaan esimerkiksi oikeusvarmuuden toteutumisen.[462] Tämä johtuu siitä, että henkilötiedoksi tulkittavan tiedon alan laajenemisen seurauksena yksittäisten toimijoiden saattaa olla vaikea arvioida, onko heidän toiminnassaan kyse henkilötietojen käsittelystä. Nykytilanteen myötä useat toimijat saattavat tulkita virheellisesti hallussaan olevien henkilötietojen olevan anonyymejä tietoja, mistä voi taas aiheutua merkittäviä tietosuojariskejä rekisteröidyille. Toisaalta henkilötiedon käsitteen laajuudesta ja epävarmuudesta voi myös seurata tilanne, jossa organisaatioista tulee ylivarovaisia siten, että kaikki tieto tulkitaan henkilötiedoksi ”varmuuden vuoksi”, joka voi vaikeuttaa tietovarantojen täysimääräistä hyödyntämistä.

Samalla henkilötiedon käsitteen määritelmän ulkopuolelle jäävän anonyymin tiedon alan on nähtävissä pienentyvän pienentymistään, kun yksilöistä kerätään yhä enemmän tietoa ja tunnistamisen mahdollistavat teknologiat kehittyvät. Esimerkiksi aikaisemmin anonyymit DNA-näytteet eivät sisältäneet henkilötietoja, sillä niiden perusteella ei ollut mahdollista tunnistaa luonnollisia henkilöitä. Tilanne kuitenkin muuttui DNA-profilointiteknologioiden kehittymisen myötä ja nykyisin DNA-

[461] Siitä huolimatta, että TSA:n tavoitteet esitetään asetuksen 1 artiklassa samanarvoisina, on henkilötietojen suoja perusoikeus toisin kuin henkilötietojen vapaa liikkuvuus. Ks. myös Julkisasiamies Bobekin ratkaisuehdotus asiassa C-40/17 FashionID, kohta 72.

[462] Oikeusvarmuudesta tarkemmin lainkäytön ennakoitavuutena ja hyväksyttävyytenä ks. Raitio 2017, s. 84–86.

näytteet tulkitaan tietosuojalainsäädännössä biometrisiksi henkilötiedoiksi.[463] Ongelmalliseksi tällaisessa tilanteessa voi muodostua henkilötiedon käsitteen dynaaminen luonne, josta johtuen anonyymeistä tiedoista voi tulla henkilötietoja jo pelkästään ajan kulumisen seurauksena, kun uusia tunnistettavuuden mahdollistavia tietoja kerätään tai julkaistaan ja tunnistamisen mahdollistavat teknologiat kehittyvät. Tämän kehityskulun seurauksena oikeuskirjallisuudessa onkin arvioitu, että lähitulevaisuudessa kaikki tieto tulee olemaan henkilötietoa, jolloin tietosuojalainsäädännöstä tulee kuvaavasti ”*the law of everything*”.[464] Se, että kaikki tieto tulkittaisiin henkilötiedoksi ei liene toivottavaa, sillä tietosuojasääntöjen ulottaminen kaiken tiedon käsittelyyn lisäisi rekisterinpitäjien henkilötietojen käsittelyyn liittyviä velvoitteita merkittävästi, muttei välttämättä tosiasiallisesti parantaisi henkilötietojen suojaa. Tämä johtuu siitä, että näin kaikenkattavan järjestelmän tehokas viranomaisvalvonta voisi olla hyvin vaikeaa, jolloin koko järjestelmän uskottavuus saattaisi horjua.

Tällä hetkellä keskeisimmät henkilötiedon käsitteen tulkitsijat EU-oikeudessa ovat tietosuojatyöryhmän korvannut EDPB ja EU-oikeuden ylimpänä tulkitsijana toimiva EUT. Uutta tulkintaa henkilötiedon käsitteen sisällön osalta ei kuitenkaan vaikuta olevan tulossa lähiaikoina, sillä EDPB:n 2019-2020 työohjelmassa[465] ei ole mainintaa henkilötiedon käsitteen konseptin tarkastelusta, ja EUT:ssa ei ole tällä hetkellä vireillä tapauksia, joissa olisi kyse henkilötiedon käsitteen yksityiskohtaisesta tulkinnasta.[466] Näin ollen henkilötiedon käsitteen nykyinen erittäin laaja tulkintaa on käytännössä tällä hetkellä vallitseva oikeustila.

Henkilötiedon käsitteen dynaamista luonnetta tarkastellessa henkilötietojen anonymisointi asettuu hieman haastavaan tilanteeseen. Tämä johtuu ennen kaikkea siitä, että TSA:n mukainen anonymisointi edellyttää prosessilta peruuttamattomuutta, josta ei kuitenkaan ole käytännössä minkäänlaisia takeita, sillä mistä tahansa tiedosta voi rekisterinpitäjästä riippumattomista seikoista johtuen tulla myöhemmin henkilötietoa. Näin ollen tiedot anonymisoineella rekisterinpitäjällä ei vaikuta olevan juurikaan mahdollisuuksia varmistua siitä, että tiedot ovat tosiasiallisesti anonymisoitu tietosuojalainsäädännön vaatimukset täyttävällä tavalla siten, että tietojen anonymiteetti säilyy myös tulevaisuudessa tietojen säilytyksen aikana.

[463] TSA 34 johdantokappale; TSA 4(1)(14).
[464] Purtova 2018, s. 41; van Der Sloot 2014, s. 309.
[465] EDPB Work Program 2019/2020.
[466] Tarkastettu 13.8.2019 EUT:n virallisilla internet-sivuilla (curia.europa.eu) olevalla hakulomakkeella valitsemalla ”*cases pending*” ja kirjoittamalla hakukenttään ”*concept of 'personal data'*”.

Tämän valossa haasteelliseksi kysymykseksi muodostuu, miten tulee tietosuojaoikeudellisesti arvioida tilannetta, jossa rekisterinpitäjä on anonymisoinut henkilötiedot nykytekniikan näkökulmasta riittävällä tavalla, mutta teknologian kehityksen seurauksena anonymisointi on mahdollista purkaa esimerkiksi viiden vuoden kuluttua.[467] Tällaisessa tapauksessa rekisterinpitäjä on nykyhetkellä toiminut tietosuojalainsäädännön vaatimukset täyttävällä tavalla, jos yksittäisiä henkilöitä ei ole tosiasiallisesti mahdollista tunnistaa tiedoista.[468] Yksinkertainen ratkaisu on, että rekisterinpitäjän tulee tekniikan kehityksen myötä palautuneen tunnistettavuuden seurauksena alkaa jälleen käsitellä kyseisiä tietoja henkilötietoina tietosuojalainsäädännön edellyttämällä tavalla. Monimutkaisempi tilanne sen sijaan on, jos käy ilmi, että henkilötietojen anonymisointi oli mahdollista purkaa jo anonymisointiprosessin aikana olemassa olevan tekniikan avulla. Tällöin kyseiset tiedot eivät itse asiassa ole koskaan olleet anonymisoituja tietoja, vaan pikemminkin ainoastaan pseudonymisoituja henkilötietoja, joiden käsittelyyn liittyviä tietosuojavelvoitteita rekisterinpitäjä ei ole virheellisen tulkinnan seurauksena noudattanut. Tämä edelleen alleviivaa sitä, että henkilötietojen anonymisointiin liittyy lähes aina rekisterinpitäjän vastuulle jäävä jäännösriski.

Ainoana poikkeuksena edellä mainittuihin anonymisoinnin riskeihin liittyen on henkilötietojen aggregoiminen siten, että suurin osa tiedoista poistetaan, jolloin kyseisten tietojen hyödynnettävyys uusiin käyttötarkoituksiin heikkenee merkittävästi. Tällä tavalla aggregoitua tilastotietoa on käytännössä mahdollista säilyttää riskeittä ja julkaista avoimena data, sillä tämänkaltaisista tilastotiedoista ei ole mahdollista tunnistaa yksittäisiä henkilöitä. Toisaalta on huomioitava, että uusia anonymisointitekniikoita kehitetään jatkuvasti ja muun muassa differentiaalisen yksityisyyden menetelmää pidetään tällä hetkellä lupaavana ratkaisuna, jonka avulla voisi olla mahdollista turvata sekä luonnollisten henkilöiden henkilötietojen suoja että tietojen hyödynnettävyys data-analytiikkaan.[469]

Henkilötietojen anonymisoinnin ohella ristiriitaiseen asemaan asettuu myös toukokuun 2019 lopussa voimaantullut asetus muiden kuin henkilötietojen vapaan liikkuvuuden kehyksestä Euroopan unionissa.[470] Asetuksella säännellään säännöistä, joiden nojalla tietosuojalainsäädännön näkökulmasta

[467] Tästä huolimatta eurooppalainen lainsäätäjä vaikuttaa luottavan henkilötietojen anonymisoinnin konseptiin, sillä muun muassa ePrivacy-asetus heinäkuun 2019 luonnos, 17aa johdantokappaleessa on vaatimus ”to genuinely anonymise the result before sharing the analysis with third parties”.

[468] Näin ollen rekisterinpitäjälle ei voitane määrätä TSA:n mukaisia sanktioita, jos anonymisointi on toteutettu nykytekniikan valossa riittävällä tavalla, sillä tällöin tiedot ovat nykyhetkellä tosiasiallisesti anonyymejä tietoja. On kuitenkin aiheellista todeta, että anonymisoitujen tietojen avoin julkaiseminen ei ole edellä mainituista syistä suositeltavaa, sillä tällöin tietojen säilytysaika ei ole enää rekisterinpitäjän määrättävissä, jolloin milloin tahansa tulevaisuudessa kehitetyn teknologian seurauksena anonymisoiduista tiedoista voi tulla henkilötietoja.

[469] Differentiaalisesta yksityisyydestä tarkemmin, ks. esim. WP 216, s. 15–16.

[470] Euroopan parlamentin ja neuvoston asetus (EU) 2018/1807 muiden kuin henkilötietojen vapaan liikkuvuuden kehyksestä Euroopan unionissa (Regulation on the free flow of non-personal data).

anonyymejä tietoja voidaan siirtää EU:ssa.[471] TSA:n mukaisen henkilötiedon käsitteen dynaamisesta luonteesta johtuen vaikuttaa kuitenkin siltä, että anonyymien tietojen liikkuvuutta koskevan asetuksen mukaista toimintaa on vaikea toteuttaa ilman henkilötietojen suojan vaarantumista, sillä toisaalle lähetetyistä anonyymeistä tiedoista voi tulla myöhemmin henkilötietoja esimerkiksi tietojen yhdistelyn seurauksena.

Tietosuoja-asetusta on sovellettu nyt hieman yli vuoden ja mitä todennäköisimmin tietosuojalainsäädännön kannalta perustavanlaatuista henkilötiedon käsitteen määritelmää ei tulla muuttamaan seuraavien vuosien aikana. Näin ollen vaikuttaa siltä, että ensisijainen keino selkeyttää tällä hetkellä perustellusti epäselvää ja vaikeasti hahmotettavaa henkilötiedon käsitettä, joka kuitenkin määrittää koko tietosuojalainsäädännön soveltumisen, on EUT:n ratkaisu, jonka perusteluissa tarkasteltaisiin yksityiskohtaisesti henkilötiedon käsitettä kaikkien sen osatekijöiden valossa. Onkin todennäköistä, että unionin tuomioistuimen seuraava henkilötiedon käsitettä koskeva ratkaisu tulee määrittämään tietosuojalainsäädännön soveltamisalan suunnan lähitulevaisuudessa. Toisaalta myöskään EDPB:n rooli ei ole merkityksetön, sillä sen lausunnot vaikuttavat *de facto* tietosuojalainsäädännön tulkitaan EU:ssa, joten kattava uusi ohjeistus joko henkilötiedon käsitteestä tai niiden anonymisoinnista olisi myös omiaan selkeyttämään henkilötiedon käsitteen tulkintaa EU-oikeudessa. Tästä huolimatta on huomioitava, että EDPB on rooliltaan neuvoa-antava viranomainen, jonka lausuntoihin EUT ei ole sidottu.[472]

Edellä mainittujen henkilötiedon käsitteen tulkintaan liittyvien haasteiden perusteella on mahdollista kritisoida koko tietosuojalainsäädännön perusteita, sillä henkilötiedon käsite muodostaa teoriassa tietosuojalainsäädännön ytimen. Kaksijakoisuuteen perustuva järjestelmä, jossa tieto joko on henkilötietoa, jolloin tietosuojasääntöjä sovelletaan, tai se ei ole henkilötietoa, jolloin käsittelylle ei ole minkäänlaisia rajoituksia, saattaa olla perusteiltaan liian yksinkertainen verkottuneessa yhteiskunnassa.[473] Tietosuojalainsäädännön henkilötiedon käsitteeseen perustuva mustavalkoinen soveltamisalan määräytyminen on pikemminkin nähtävissä eurooppalaisen lainsäätäjän yrityksenä yksinkertais-

[471] Asetuksen 2018/1807 soveltamisalaa koskevassa 2 artiklassa säädetään, että ”*Tätä asetusta sovelletaan sellaiseen muiden sähköisten tietojen kuin henkilötietojen käsittelyyn unionissa*” ja asetuksen määritelmiä koskevassa 3 artiklassa tiedoilla tarkoitetaan ”*muita tietoja kuin asetuksen (EU) 2016/679 4 artiklan 1 alakohdassa määriteltyjä henkilötietoja*”.
[472] On myös kuvaavaa, että EUT:lla ei ole ollut tapana viitata tietosuojatyöryhmän lausuntoihin ja kannanottoihin millään tapaa ratkaisuissaan.
[473] Purtova 2018, s. 78–80; Hintze 2018, 89; Koops 2014, s. 258.

taa monimutkaista tietosuojaan liittyvää maailmaa, jossa erilaisten tietojen käsittelyllä on tosiasiallisesti erilaisia vaikutuksia rekisteröityjen perusoikeuksien kannalta.[474] Toisaalta voidaan myös perustellusti sanoa, että tietojen käsittelyyn liittyvät tilanteet, tekniikat ja ympäristöt ovat monimutkaistuneet huomattavasti sitten 1980-luvun, jolloin henkilötiedon käsitettä kehitettiin.

On kuitenkin toisaalta huomioitava, että tietosuojalainsäädännön osalta *law in books* ja *law in action* ovat eronneet toisistaan merkittävästi. Henkilötiedon käsitteen ala oli todella laaja jo henkilötietodirektiivin aikana, mutta henkilötietojen suojaa koskevan lainsäädännön merkitys korostui huomattavasti vasta TSA:n myötä.[475] Siitä syystä, että organisaatiot eivät aikaisemmin olleet erityisen tietoisia omista tietosuojavelvoitteistaan, on laaja henkilötiedon käsite ja tehokas tietosuojasääntöjen soveltaminen EU:ssa hyvinkin perusteltua. Tässä merkittävä rooli on kansallisilla tietosuojaviranomaisilla, joilla on muun muassa TSA:n 83 artiklan nojalla oikeus määrätä sanktioita organisaatioille, jotka eivät noudata tietosuojasääntöjä henkilötietojen käsittelyssä.[476] Sanktioiden määrääminen voi tulla kyseeseen esimerkiksi silloin, jos rekisterinpitäjä laiminlyö tietosuoja-asetuksen velvoitteita tai pitää hallussaan olevia henkilötietoja virheellisen tulkinnan seurauksena anonyymeinä tietoina, eikä tästä johtuen noudata tietosuojasääntöjä lainkaan näiden tietojen osalta. Toisaalta tarkasteltaessa tilannetta sen valossa, että kaikki tieto voi tulevaisuudessa olla henkilötietoa, on vaarana, että tietosuojasääntely ja tietojenkäsittelyn todellisuus tulevat eroamaan toisistaan vieläkin merkittävämmin, sillä *kaikkea* tietojenkäsittelyä ei ole viranomaistenkaan tarkoituksenmukaista valvoa.

Eurooppalaisen tietosuojalainsäädännön merkitys on kasvanut läpi 2010-luvun Lissabonin sopimuksella itsenäiseksi perusoikeudeksi kirjatun henkilötietojen suojan myötä. Tämä kehitys kulminoituu tietosuoja-asetukseen, jonka seurauksena EU:n kansalaiset ovat nykyisin tietoisempia omista tietosuojaoikeuksistaan kuin koskaan aikaisemmin. TSA:n tavoitteina ovat henkilötietojen suoja ja henkilötietojen vapaa liikkuvuus, joista aiempi on kuitenkin dominoinut niin kansallisten tietosuojaviranomaisten, tietosuojatyöryhmän kuin EUT:n tietosuojalainsäädännön soveltamiseen liittyviä tulkintoja. Osin henkilötietojen tehokkaan suojan merkityksen korostumisen seurauksena myös itse henkilötiedon käsite, se mitä tietosuojalainsäädännöllä suojellaan, on laajentunut kattamaan mitä moninaisempia tietoja, joiden käsittelyllä voi olla vaikutuksia luonnollisten henkilöiden perusoikeuksien toteutumisen kannalta. Näin ollen henkilötiedon käsitteestä on muodostunut avoin, dynaaminen ja

[474] Koops 2014, s. 250.
[475] Euroopan komissio 2019, s. 18. Komission 24.7.2019 julkaistun tiedonannon mukaan TSA:n soveltamisen ensimmäinen vuosi on ollut kokonaisuutena positiivinen ja tietosuoja otetaan nykyisin vakavammin kuin koskaan aikaisemmin.
[476] TSA:n 83 artiklan mukaisista hallinnollisista sanktioista tarkemmin, ks. esim. Koillinen 2016, s. 570–572.

teknologianeutraali konsepti, jonka tämänhetkisessä eurooppalaisessa tietosuojalainsäädännössä omaksutun tulkinnan mukaan kaikki tieto joko on, tai voi olla henkilötietoa.